

UNIVERSIDADE FEDERAL DO PARANÁ

DANIELA MIRAY IGARASHI

ANÁLISE DE RISCO DE CRÉDITO POR MEIO DO MODELO CREDITRISK+: UMA
APLICAÇÃO À CARTEIRAS DE CRÉDITO

CURITIBA

2019

DANIELA MIRAY IGARASHI

ANÁLISE DE RISCO DE CRÉDITO POR MEIO DO MODELO CREDITRISK+: UMA
APLICAÇÃO À CARTEIRAS DE CRÉDITO

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, como requisito para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Luiz Carlos Matioli

CURITIBA

2019

I24a

Igarashi, Daniela Miray

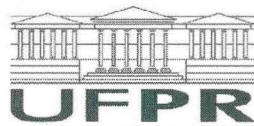
Análise de risco de crédito por meio do modelo Creditrisk+: uma aplicação à carteiras de crédito[recurso eletrônico] / Daniela Miray Igarashi , 2019.

Dissertação (mestrado) - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná.

Orientador: Prof Luiz Carlos Matioli, DSc

1. Administração de crédito. 2. Risco de crédito. I. Universidade Federal do Paraná. II. Matioli, Luiz Carlos. III. Título.

CDD 658.88



MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MÉTODOS NUMÉRICOS
EM ENGENHARIA - 40001016030P0

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **DANIELA MIRAY IGARASHI** intitulada: **ANÁLISE DE RISCO DE CRÉDITO POR MEIO DO MODELO CREDITRISK+: UMA APLICAÇÃO À CARTEIRAS DE CRÉDITO**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 14 de Fevereiro de 2019.


LUIZ CARLOS MATIOLI

Presidente da Banca Examinadora (UFPR)


ANSELMO CHAVES NETO

Avaliador Interno (UFPR)

O Professor Paulo Participou VIA SKYPE

PAULO SERGIO MARQUES DOS SANTOS

Avaliador Externo (UFPI)


GUSTAVO VALENTIM LOCH

Avaliador Interno (UFPR)

*À minha mãe,
que foi a primeira a me ensinar como ver o mundo.
Ao meu companheiro de vida e amor,
que anseia o momento de me chamar de mestre Miray.*

AGRADECIMENTOS

Os meus sinceros agradecimentos vão para meu orientador professor Dr. Luiz Carlos Matioli, pelo apoio, confiança, e paciência, depositados em mim no decorrer deste trabalho.

À minha família e amigos, pelo apoio incondicional nesta jornada. Obrigada por entenderem minha ausência nos momentos mais preciosos.

Aos meus amigos do programa, por todas as suas contribuições acadêmicas e pessoais, pelos momentos de descontração e amizade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio neste estudo.

À todos que direta ou indiretamente contribuíram, muito obrigada.

A vida é inerentemente arriscada. Existe apenas um grande risco que você deve evitar a todo custo, e esse é o risco de não fazer nada.

Denis Waitley

RESUMO

O presente trabalho apresenta o modelo CreditRisk+ como técnica de estimação do risco de crédito de duas carteiras de crédito distintas. A primeira é proveniente de uma empresa de securitização, em que a carteira é composta por devedores com contratos de financiamentos de imóveis. A segunda carteira compreende dados de crédito de empréstimos em uma instituição financeira alemã. O propósito do modelo CreditRisk+ é determinar a distribuição de probabilidade das perdas em carteiras de empréstimos ou outros instrumentos de dívida. Com a distribuição de perdas é feita estimação da perda da carteira por meio de duas métricas de risco: o Valor em Risco e o Valor em Risco Condicional, sendo a primeira uma métrica de perda potencial máxima, e a segunda uma medida de perda média excedente ao Valor em Risco. O objetivo central deste trabalho é quantificar o risco de crédito para as duas carteiras disponíveis, além de simular os efeitos dos diferentes parâmetros de entrada do modelo. Os resultados encontrados sugerem que o modelo CreditRisk+ superestima o risco da carteira de crédito quando a probabilidade de *default* dos devedores é alta, quando o número de faixas de exposição é pequena, e a volatilidade das taxas de *default* adotadas são altas.

Palavras-chaves: Risco de Crédito. Análise de Risco. CreditRisk+. *Default*. Value-at-Risk. Regressão Logística.

ABSTRACT

This dissertation presents the CreditRisk+ model as a credit risk estimation technique for two different credit portfolios. The first portfolio comes from a securitization company, where it is composed of debtors with real estate financing agreements. The second portfolio comprises loan credit data at a German financial institution. The purpose of the CreditRisk+ model is to estimate the probability distribution of losses in loan portfolios or other debt instruments. With the distribution of losses, the portfolio risk is estimated through two risk metrics: Value at Risk and Conditional Value at Risk. Value at Risk measures potential loss over a period of time, and Conditional Value at Risk is a measure of average loss exceeding Value at Risk. The main objective of this dissertation is to quantify the credit risk for the two current portfolios, as well as to simulate the effects of the different input parameters of the model. The results suggest that the CreditRisk+ model overestimates the portfolio credit risk when the default probability of the debtors is high, when the number of exposure bands is small and in cases where the volatility of the default rates are high.

Key-words: Credit Risk. Risk Analysis. CreditRisk+. Default. Value-at-Risk. Logistic Regression.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – Representação do VaR e CVaR para uma distribuição de perdas . . .	26
FIGURA 2 – Processo de Securitização	51
FIGURA 3 – Carteira 1 - Curva ROC	59
FIGURA 4 – Carteira 2 - Curva ROC	62
FIGURA 5 – Carteira 1 - Distribuição de perdas - modelo padrão	67
FIGURA 6 – Carteira 2 - Distribuição de perdas - modelo padrão	69
FIGURA 7 – Carteira 1 - Distribuição de perdas para $PD < 0,5$	70
FIGURA 8 – Carteira 1 - Distribuição de perdas para diferentes faixas de exposição (50, 100 e 300)	75
FIGURA 9 – Carteira 1 - Distribuição de perdas para diferentes faixas de exposição (100, 300 e 500)	75
FIGURA 10 – Carteira 1 - Distribuição de perdas para diferentes volatilidades da PD (0,2, 0,4 e 0,5)	78
FIGURA 11 – Carteira 1- Distribuição de perdas para diferentes volatilidades da PD (0,5, 0,6 e 0,8)	79

LISTA DE TABELAS

TABELA 1 – Exemplo 1 - Dados da carteira	43
TABELA 2 – Exemplo 2 - Dados da carteira	45
TABELA 3 – Exemplo 2 - Dados de entrada do modelo CR+	45
TABELA 4 – Exemplo 2 - Iterações do algoritmo	48
 TABELA 5 – Carteira 1 - Variáveis explicativas da carteira	 52
TABELA 6 – Carteira 2 - Variáveis explicativas da carteira	54
TABELA 7 – Carteira 2 - Codificação <i>dummy</i> para o atributo <i>Personal</i>	55
TABELA 8 – Carteira 1 - Matriz de correlação das variáveis explicativas do modelo	57
TABELA 9 – Carteira 1 - Coeficientes do modelo de regressão	58
TABELA 10 – Carteira 1 - Matriz de classificação do modelo de regressão	58
TABELA 11 – Carteira 1 - Distribuição da Probabilidade de <i>default</i>	60
TABELA 12 – Carteira 2 - Matriz de Coeficientes do modelo de regressão	61
TABELA 13 – Carteira 2 - Matriz de classificação do modelo de regressão	61
TABELA 14 – Carteira 2 - Distribuição da Probabilidade de <i>default</i>	62
 TABELA 15 – Aproximação logarítmica pela expansão de Taylor	 65
TABELA 16 – Carteira 1 - Dados da carteira de crédito	66
TABELA 17 – Carteira 1 - Perda estimada da carteira	67
TABELA 18 – Carteira 2 - Dados da carteira de crédito	68
TABELA 19 – Carteira 2 - Perda estimada da carteira	68
TABELA 20 – Carteira 1 - Dados da carteira de crédito para $PD < 0,5$	70
TABELA 21 – Carteira 1 - Perda estimada da carteira para $PD < 0,5$	70
TABELA 22 – Carteira 1 - Dados da carteira de crédito para $PD < 0,09$	71
TABELA 23 – Carteira 1 - Perda estimada da carteira para $PD < 0,09$	71
TABELA 24 – Carteira 2 - Dados da carteira de crédito para $PD < 0,5$	72
TABELA 25 – Carteira 2 - Perda estimada da carteira para $PD < 0,5$	72
TABELA 26 – Carteira 2 - Dados da carteira de crédito para $PD < 0,09$	72
TABELA 27 – Carteira 2 - Perda estimada da carteira para $PD < 0,09$	72
TABELA 28 – Carteira 1 - Perda estimada (VaR) para diferentes números de faixas	74
TABELA 29 – Carteira 1 - Perda estimada (CVaR) para diferentes números de faixas	74
TABELA 30 – Carteira 2 - Perda estimada (VaR) para diferentes números de faixas	76
TABELA 31 – Carteira 2 - Perda estimada (CVaR) para diferentes números de faixas	76
TABELA 32 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω	77
TABELA 33 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω	77
TABELA 34 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω e $PD < 0,05$	79
TABELA 35 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω e $PD < 0,05$	79

TABELA 36 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω e $PD < 0,0980$

TABELA 37 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω e $PD < 0,0980$

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS	16
1.2	ESTRUTURA DA DISSERTAÇÃO	17
2	CONCEITOS PRELIMINARES	18
2.1	CONCEITOS DA TEORIA DE PROBABILIDADES	18
2.1.1	Espaço de Probabilidade	18
2.1.2	Função Geradora de Probabilidade	19
2.1.3	Distribuição de Poisson	19
2.1.4	Função gama	19
2.1.5	Distribuição Gama	20
2.1.6	Distribuição Composta	20
2.2	TIPOS DE RISCOS FINANCEIROS	20
2.3	MODELOS DE <i>CREDIT RATING</i> E AGÊNCIAS DE CLASSIFICAÇÃO DE RISCO	21
2.4	MODELOS DE <i>CREDIT SCORING</i> E REGRESSÃO LOGÍSTICA . . .	22
2.5	VALOR EM RISCO	23
2.6	VALOR EM RISCO CONDICIONAL	25
2.7	MODELOS DE RISCO DE CARTEIRAS DE CRÉDITO	26
2.7.1	CreditMetrics	26
2.7.2	CreditPortfolioView	27
2.7.3	KMV	28
2.7.4	CreditRisk+	28
3	O MODELO CREDITRISK+	30
3.1	CARACTERÍSTICAS E SUPOSIÇÕES	30
3.2	COMPONENTES DE ENTRADA DO MODELO CREDITRISK+	31
3.3	DISTRIBUIÇÃO DE PERDA	33
3.4	FORMA RECURSIVA	37
3.5	EXEMPLOS DA APLICAÇÃO DO MODELO CREDITRISK+	42
3.5.1	Ilustração 1	42
3.5.2	Ilustração 2	44
4	DADOS E MÉTODOS	50
4.1	CARTEIRA DE CRÉDITO DE SECURITIZAÇÃO	50

4.2	CARTEIRA DE CRÉDITO ALEMÃ	53
4.3	ESTIMAÇÃO DA PROBABILIDADE DE <i>DEFAULT</i>	55
4.3.1	Carteira 1	56
4.3.2	Carteira 2	60
4.4	ESTIMAÇÃO DA VOLATILIDADE DA TAXA DE <i>DEFAULT</i>	62
5	RESULTADOS NUMÉRICOS	64
5.1	EFEITO DA PROBABILIDADE DE <i>DEFAULT</i>	65
5.1.1	Estimativa de Perda	65
5.1.2	Probabilidade de <i>default</i> abaixo de 50% e 9%	69
5.2	EFEITO DO NÚMERO DE FAIXAS E TAMANHO DA UNIDADE DE PERDAS	73
5.3	EFEITO DA VOLATILIDADE DA TAXA DE <i>DEFAULT</i>	76
5.3.1	Probabilidade de <i>default</i> abaixo de 50% e 9%	78
6	CONCLUSÃO	81
	REFERÊNCIAS	83

1 INTRODUÇÃO

Durante a história dos mercados financeiros, algumas crises e quedas de instituições financeiras estiveram relacionadas diretamente aos eventos simultâneos de inadimplência. Um exemplo deste fenômeno é a crise do *subprime* em 2008, o que mostra como a gestão e avaliação de risco são essenciais para a saúde do sistema financeiro, e consequentemente, um desafio diário para os gestores da área.

O risco na área de finanças pode ser definido como uma medida da incerteza dos resultados de um evento. Alguns destes eventos inesperados podem ser a mudança nas taxas de juros ou a inadimplência de obrigações financeiras, por exemplo (JORION, 2001). Quando um evento é previsível, é possível planejar de acordo e tomar decisões para conter possíveis perdas, sendo assim, o risco financeiro surge de resultados que ainda ocorrerão e não são passíveis de previsão.

Como o risco financeiro é um fator permanente para empresas em funcionamento, é tarefa da organização administrar e reduzir possíveis ameaças ao seu negócio. A necessidade da gestão de riscos é ressaltada quando considera-se o seu potencial em causar perdas e danos ao equilíbrio patrimonial e, em casos extremos, ser responsável pela queda de toda a instituição.

Neste sentido, todas as empresas estão expostas aos diferentes tipos de riscos, dentre os quais pode-se citar o risco de mercado, risco de liquidez, risco operacional, risco legal, risco de crédito e outros (JORION, 2001). No contexto do presente trabalho, tem destaque o risco de crédito.

Para Caouette, Altman e Narayanan (2008), uma das definições para o conceito de crédito é uma expectativa de recebimento de recursos por um certo período de tempo. O mesmo pode ser entendido também como o ato de colocar a disposição de um cliente um certo valor monetário, em forma de empréstimo ou financiamento, com o compromisso de que haverá pagamento futuro (BRITO; ASSAF NETO, 2008). Dessa forma, o risco de crédito pode ser entendido como o evento em que essas obrigações não são cumpridas, provocando perdas.

O risco de inadimplência é, então, a possibilidade de que uma das partes em um contrato financeiro não cumpra os compromissos contratuais estabelecidos anteriormente. Se isso vier a ocorrer, determina-se que a parte é inadimplente ou que um evento de *default* acontece. O termo *default* é utilizado neste contexto para indicar o descumprimento de obrigações entre o devedor e credor, podendo ser traduzido também como a expressão “calote”.

De forma geral, o risco de crédito é o risco associado a qualquer tipo de operação

de crédito. Dependendo do contexto analisado, o risco de crédito pode ser vinculado à alterações na qualidade de crédito, variações de *spreads*¹ de crédito e o próprio evento de *default* (BIELECKI; RUTKOWSKI, 2004).

Caouette, Altman e Narayanan (2008) afirmam que uma parcela de risco de crédito sempre surge quando um indivíduo retira um produto ou serviço sem pagar imediatamente por ele. Empresas de telefonia e concessionários de energia elétrica convivem com o risco de crédito de todos os seus clientes. Emissoras de cartão de crédito assumem esse risco com todos os titulares dos cartões, assim como as instituições financeiras o fazem ao conceder empréstimos. No setor de vendas, lojistas em todos os níveis oferecem aos seus clientes termos de pagamentos à prazo. Em todas estas situações, as empresas aceitam o risco de crédito.

Dessa forma, a gestão de risco de crédito é um processo essencial para saúde de qualquer empresa e instituição financeira. Sendo assim, abordagens para a administração do risco financeiro são implementadas em diversos setores. Formalmente, pode-se definir a gestão de risco como o processo de identificar e analisar a incerteza envolvida nas decisões das operações financeiras. Sendo que, na prática, a gestão de risco ocorre quando um gestor de capital tenta quantificar a perda potencial em um dado investimento, analisando quais decisões devem ser tomadas ou não para a mitigação do risco.

Considerando que raramente existem operações livre de riscos, a administração de riscos se refere essencialmente à como as empresas selecionam o tipo e o nível de risco que é apropriado assumirem (CROUHY; GALAI; MARK, 2005). Dessa forma, todas as decisões e ações tomadas são orientadas considerando a tolerância ao risco do gestor e o propósito do investimento sendo realizado.

A gestão de risco surgiu naturalmente como uma resposta ao aumento da incerteza nos mercados financeiros globais. Apesar da literatura de gestão de risco ter desenvolvimentos desde a metade do século passado, técnicas e modelos aplicados especificamente à quantificação do risco de crédito foram desenvolvidas na década de 90. Os principais modelos que se destacaram nesta tarefa foram o Creditmetrics, KMV, Credit Portfolio View, e o modelo CreditRisk+ (CR+).

Todos estes modelos de risco de crédito seguem uma abordagem semelhante na gestão do risco, no sentido que o principal objetivo é calcular a distribuição de probabilidade de perdas em um determinado horizonte de tempo quando há *default* (BIELECKI; RUTKOWSKI, 2004). A distribuição de perda é um elemento importante na gestão da carteira, pois é utilizada para quantificação do risco por meio de métricas como o Valor em Risco.

¹ O conceito de *spread* está geralmente associado à diferença entre dois preços, de compra e venda, para algum título ou rendimento. Neste contexto, o *spread* é a diferença entre a taxa cobrada pelo empréstimo e a taxa de captação da instituição financeira.

Os três primeiros modelos citados são modelos de marcação a mercado, ou seja, as mudanças nos níveis de risco dos devedores refletem as variações de valor de mercado. São modelos que exigem de dados de entrada como matrizes de migração de crédito, matrizes de *spread*, do valor de mercado da carteira de empréstimos e diversos outros. Por necessitarem de informações que por vezes não estão disponíveis no mercado financeiro brasileiro, estes modelos possuem características que dificultam a sua aplicação no contexto atual. A seção 2.7 apresenta uma descrição breve de cada um dos modelos.

Por outro lado, o modelo CreditRisk+ é classificado como um modelo de *default*, ou seja, em vez de depender de alterações na qualidade de crédito do devedor ou do mercado para mensurar o risco, este modelo se concentra no evento de *default* de cada devedor para a estimação da distribuição de perdas. O modelo CR+ utiliza técnicas estatísticas, em que não são feitas suposições sobre as causas da inadimplência, e a distribuição de eventos de *default* é aproximada pela distribuição de Poisson. Estas e outras suposições permitiram o desenvolvimento de uma expressão analítica para o cálculo da distribuição de perdas, o que isenta o modelo da necessidade de métodos de simulações de Monte Carlo para a estimação da distribuição de probabilidades.

Por estas e outras características, o modelo CR+ é atrativo para os gestores em mercados como o brasileiro. Segundo Avesani et al. (2014) e Gundlach e Lehrbass (2004), algumas razões para a popularidade do modelo e justificativas para sua escolha são:

- a) Exige uma quantidade limitada de dados de entrada e suposições teóricas;
- b) Utiliza os mesmos dados de entrada que são exigidos pelo Acordo de Basiléia II na abordagem IRB²;
- c) Fornece uma solução analítica eficiente para determinar a distribuição de perdas, evitando a utilização de simulações de Monte Carlo.

Assim, o modelo CreditRisk+ será utilizado para quantificação do risco de crédito no presente trabalho. A aplicação do modelo CreditRisk+ será feita para duas carteiras de crédito. A primeira carteira compreende dados cedidos por uma empresa de securitização, em que os devedores possuem contratos de financiamentos de imóveis em um condomínio. A segunda carteira é proveniente de uma base de dados, em que os devedores possuem empréstimos de uma instituição financeira alemã.

1.1 OBJETIVOS

Considerando os aspectos apresentados, o principal objetivo deste trabalho é quantificar o risco de crédito ao qual as carteiras em análise estão expostas por meio do

² A abordagem IRB (*Internal ratings based*) se refere ao conjunto de técnicas de estimação de risco de crédito proposta pelo Acordo de Basiléia II para instituições bancárias. O Acordo de Basiléia II foi proposto pelo Comitê de Supervisão Bancária da Basiléia (*Basel Committee on Banking Supervision*), de forma a trazer práticas e políticas de gestão para fortalecer as instituições financeiras.

modelo CreditRisk+. Isto é feito por meio da distribuição de probabilidade de perdas encontrada pelo modelo, e o risco poderá ser estimado pelas métricas Valor em Risco e Valor em Risco Condicional.

Este trabalho também tem o intuito de avaliar a sensibilidade do modelo aos diferentes parâmetros de entrada. Para tanto, a perda será estimada para diferentes valores de cada parâmetro a fim de observar quais são os impactos de cada dado de entrada no resultado da estimação do risco de crédito da carteira.

1.2 ESTRUTURA DA DISSERTAÇÃO

O presente trabalho está organizado em 6 capítulos. No capítulo 2, alguns conceitos base necessários são apresentados como forma de preparação para a teoria desenvolvida nos capítulos subsequentes. O capítulo 3 é dedicado ao modelo CreditRisk+, em que são apresentados desde suas características e suposições iniciais, detalhes dos dados de entrada, até o desenvolvimento da teoria que resulta na sua expressão recursiva. O capítulo 4 mostra os dados utilizados no presente trabalho, ou seja, as carteiras de crédito. Neste capítulo, é discutido também como os parâmetros de entrada do modelo CreditRisk+ são estimados e considerados. No capítulo 5, tem-se os resultados das estimações de risco realizadas para as carteiras de crédito. Diferentes valores são adotados para cada dado de entrada do modelo, de forma a avaliar os efeitos de cada parâmetro na estimação de perda resultante. Finalmente, o capítulo 6 apresenta as considerações finais.

2 CONCEITOS PRELIMINARES

O presente capítulo tem o intuito de apresentar os conceitos preliminares necessários para o entendimento da teoria do modelo CreditRisk+ e demais desenvolvimentos realizados em capítulos posteriores.

2.1 CONCEITOS DA TEORIA DE PROBABILIDADES

Esta seção apresenta brevemente algumas noções gerais e notações da teoria de probabilidades, servindo como base teórica do desenvolvimento dos capítulos posteriores. As referências para esta seção são Gundlach e Lehrbass (2004), Johnson, Kotz e Kemp (1992) e James (1996).

2.1.1 Espaço de Probabilidade

Inicialmente, considera-se um modelo probabilístico para a realização de experimentos. Definindo Ω como o conjunto de resultados possíveis, entende-se que um resultado possível é todo resultado elementar e indivisível do experimento. Dessa forma, diz-se que o conjunto Ω é o espaço amostral do experimento.

Dado um espaço amostral Ω , o subconjunto $F \subset \Omega$ é chamado de evento. Define-se também o conjunto Ω como evento certo e o conjunto \emptyset como evento impossível. Seja F um subconjunto do espaço amostral Ω e se os resultados do evento F são equiprováveis, uma definição clássica para a medida da probabilidade (\mathbb{P}) de ocorrência do evento F pode ser dada por

$$\mathbb{P}(F) = \frac{\#F}{\#\Omega} \quad (2.1)$$

em que $\#F$ representa o número de resultados favoráveis de F e $\#\Omega$ representa o número de resultados possíveis. Quando uma medida de probabilidade é atribuída ao evento F , tem-se um evento aleatório.

Supondo que a classe desses eventos aleatórios possua certas propriedades, sendo necessárias para o desenvolvimento da teoria de probabilidades, define-se \mathcal{F} como a classe de eventos aleatórios, com as seguintes propriedades:

- a) $\Omega \in \mathcal{F}$
- b) Se $F \in \mathcal{F}$, então $F^c \in \mathcal{F}$
- c) Se $F_n \in \mathcal{F}$ para $n = 1, 2, \dots$, então $\bigcup_{n=1}^{\infty} F_n \in \mathcal{F}$

Dessa forma, se Ω é um conjunto não vazio, a classe \mathcal{F} de subconjuntos de Ω que satisfaz os itens anteriores é chamada de σ -álgebra.

Finalmente, considera-se um espaço de probabilidade como a tripla $(\Omega, \mathcal{F}, \mathbb{P})$, conforme definido anteriormente, sendo este espaço escolhido de forma a manter todas as variáveis mensuráveis.

No decorrer desta dissertação, denota-se também por \mathbb{E} a esperança das variáveis aleatórias, assim como var denota a variância em relação ao \mathbb{P} . Para uma variável aleatória Y no espaço de probabilidade, a sua função densidade de probabilidade é escrita como f_Y , se existir. Se uma variável aleatória Y segue uma distribuição conhecida f , representa-se por $Y \sim f$.

2.1.2 Função Geradora de Probabilidade

Dada a variável aleatória Y que possui apenas valores inteiros não negativos, sua distribuição pode ser descrita pela função geradora de probabilidade (FGP) G_Y definida como

$$G_Y(z) = \sum_{n \geq 0} \mathbb{P}[Y = n] z^n = \mathbb{E}[Z^Y] \quad (2.2)$$

para todo $z \in \mathbb{R}$ em que a sequência (2.2) converge e $n = 0, 1, \dots$. Denota-se também $\pi(n) = \mathbb{P}[Y = n]$. A variável z não possui interpretação por si própria. Tem-se também que se a sequência π é limitada, então (2.2) converge para $|z| \leq 1$.

A FGP contém as informações necessárias para calcular a distribuição de probabilidade, sendo esta obtida por meio da n -ésima derivada

$$\pi(n) = \mathbb{P}[Y = n] = \frac{G_Y^{(n)}(0)}{n!}. \quad (2.3)$$

2.1.3 Distribuição de Poisson

Uma variável aleatória Y segue uma distribuição de Poisson com parâmetro θ se

$$\mathbb{P}[Y = y] = \frac{e^{-\theta} \theta^y}{y!} \quad (2.4)$$

em que $y = 0, 1, \dots$, e $\theta > 0$.

A FGP desta distribuição é dada por

$$G(z) = e^{\theta(z-1)}. \quad (2.5)$$

2.1.4 Função gama

A função gama, denotada por Γ , é definida por

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0. \quad (2.6)$$

Quando $x > 0$ e n é um inteiro positivo, a equação (2.6) tem significado

$$\Gamma(n+1) = n!. \quad (2.7)$$

2.1.5 Distribuição Gama

O modelo CreditRisk+ é construído sob a premissa de que diferentes setores podem influenciar na probabilidade de *default* dos devedores da carteira de crédito. Nos desenvolvimentos do modelo, supõe-se que o comportamento dos diferentes setores pode ser explicado por variáveis aleatórias com distribuição gama.

Seja a variável Y com distribuição gama com parâmetros $\alpha > 0$ e $\beta > 0$, denotando-se por $Y \sim Gama(\alpha, \beta)$. Sua função densidade de probabilidade é dada por

$$f_Y^{(\alpha, \beta)}(y) = \frac{y^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-y/\beta}, \quad (2.8)$$

para $y > 0$ e a função gama Γ . A esperança e variância da variável Y são dadas por

$$\begin{aligned} \mathbb{E}[Y] &= \alpha\beta \\ \text{var}[Y] &= \alpha\beta^2. \end{aligned} \quad (2.9)$$

2.1.6 Distribuição Composta

Uma distribuição de probabilidade composta é uma distribuição obtida por meio da suposição de que uma variável aleatória é distribuída por um parâmetro dependente, sendo este parâmetro também uma variável aleatória com distribuição conhecida.

A expressão (2.10) representa a função densidade de probabilidade de uma variável aleatória z . Esta variável é uma distribuição composta de uma distribuição F parametrizada por θ que é, por sua vez, distribuído por uma outra distribuição G .

$$p_X(z) = \int p_F(z|\theta) p_G(\theta) d\theta. \quad (2.10)$$

2.2 TIPOS DE RISCOS FINANCEIROS

O conceito de risco é definido em relação aos eventos esperados e inesperados, sendo o risco uma medida associada à incerteza e volatilidade dos resultados destes eventos. Entretanto, o risco pode ser dividido em categorias de acordo com a natureza da exposição ao risco de cada empresa. Portanto, as empresas podem ser expostas aos diferentes tipos de riscos, dentre os quais pode-se citar o risco de mercado, risco de liquidez, risco operacional, risco legal, risco de crédito e outros.

Desse modo, a presente seção faz um sumário de alguns tipos de riscos financeiros. Estas descrições são feitas baseadas nos trabalhos de Jorion (2001) e Crouhy, Galai e Mark (2001).

O risco de mercado é definido como o potencial de perda decorrente de mudanças nos preços ou parâmetros do mercado. O risco de mercado pode surgir a partir de mudanças nas taxas de juros, taxas de câmbio ou preços de ações, dentre outros.

O risco de crédito surge na possibilidade de uma contraparte ser incapaz de cumprir suas obrigações contratuais, entrando em estado de *default* e causando perdas para a instituição. O risco de crédito pode também ser definido como a possível perda no valor de mercado devido à ocorrência de eventos de crédito. Um evento de crédito ocorre quando há uma mudança na capacidade da contraparte de cumprir suas obrigações, assim, mudanças nos preços da dívida devido a mudanças no *rating* de crédito ou na percepção de inadimplência do mercado, também podem ser vistas como risco de crédito.

O risco de liquidez é relacionado ao conceito de liquidez que, no contexto financeiro, é definido como a velocidade com que um dado patrimônio pode ser convertido em valores monetários. O risco de liquidez pode ser dividido em duas classes gerais: o risco de liquidez de mercado e o risco de liquidez de fluxo de caixa.

O primeiro está associado à incapacidade de executar uma transação ao preço de mercado porque não há, temporariamente, demanda por este produto. Se a transação não puder ser adiada, sua realização irá causar perdas para a instituição, pois os preços precisarão ser menores para que a liquidação ocorra. Já o risco de liquidez de fluxo de caixa está relacionado ao montante de capital que uma instituição deve ter para honrar todos os seus compromissos assumidos. A perda ocorre quando a instituição é incapaz de captar o caixa necessário para atender todos estes compromissos, sendo por vezes obrigada a liquidar ativos precocemente, o que diminui sua rentabilidade.

O risco operacional pode ser definido como as perdas financeiras resultantes de erros ou acidentes humanos e técnicos. Dentre os fatores que contribuem para o risco operacional, podem ser citadas as fraudes, falhas no gerenciamento e procedimentos, e controles feitos de forma inadequada. Já os erros técnicos podem ser causados por falhas nas informações, nos processamento de transações, e nos sistemas de liquidação, por exemplo.

O risco legal surge de decisões relacionadas à lei, ocasionado por multas ou indenizações em processos judiciais. Este tipo de risco pode ser também relacionado ao risco de crédito, já que uma contraparte em *default* ou um investidor que perdeu dinheiro podem tentar encontrar bases legais para invalidar a operação e, dessa forma, não cumprir com suas obrigações. Outro aspecto deste tipo de risco é o de natureza regulatória, em que mudanças em leis, por exemplo, podem causar um impacto negativo em empresas e instituições.

2.3 MODELOS DE *CREDIT RATING* E AGÊNCIAS DE CLASSIFICAÇÃO DE RISCO

As agências de classificação de risco são empresas especializadas em avaliar a qualidade de crédito dos produtos ou ativos financeiros emitidos por empresas ou governos. Esta classificação pode ser feita na forma de categorias ordenadas, como classificações A,

B, ou C, ou na forma de pontuações, gerando o que é conhecido como *score* de crédito (CAOINETTE; ALTMAN; NARAYANAN, 2008).

Neste contexto, uma classificação de crédito ou *rating* é a opinião de uma agência de classificação sobre a credibilidade geral de um devedor em relação a um determinado produto ou obrigação financeira, sendo baseada em certos fatores de risco (CROUHY; GALAI; MARK, 2001). As principais empresas que fornecem estas classificações em nível global são as agências Moody's e Standard & Poors's (S&P). No Brasil, a principal fornecedora de classificações de crédito é a SERASA.

Estes *ratings* são então produzidos por meio de modelos de *credit rating*. Segundo Crouhy, Galai e Mark (2001), o processo de classificação inclui análises quantitativas, qualitativas e legais. A análise quantitativa tem relação com a situação financeira da contraparte avaliada e baseia-se nos relatórios de demonstrações contábeis da empresa. A análise qualitativa é associada a fatores diversos como, por exemplo, a qualidade da gestão da empresa avaliada, a competitividade do setor de atuação, o crescimento esperado da indústria e as vulnerabilidades a mudanças do mercado. Ou seja, um elemento de julgamento é presente nestas avaliações.

2.4 MODELOS DE *CREDIT SCORING* E REGRESSÃO LOGÍSTICA

Os modelos de *Credit Scoring* são modelos quantitativos comumente utilizados por instituições financeiras de modo a avaliar e mensurar o risco de seus devedores numa operação de crédito. Estes métodos utilizam as características do solicitante para a atribuição do risco, sendo considerados modelos de classificação de risco, pois visam a atribuição de categorias aos devedores.

Da mesma forma como nos modelos de *credit rating*, os modelos de *credit scoring* oferecem uma classificação ou pontuação de risco para cada devedor. Neste sentido, o propósito destes métodos geralmente é discriminar entre bons e maus pagadores, avaliando o nível de risco de cada indivíduo.

Segundo Bolton (2009), antes da aplicação de métodos formais no setor bancário, as decisões eram tomadas baseadas no julgamento. O gerente do banco avaliaria a credibilidade de um indivíduo com base em seu conhecimento pessoal, o que representava várias deficiências. Dentre estas pode-se citar o fato do julgamento ser subjetivo, não replicável e não confiável, além de ser ineficiente com um grande número de clientes.

Com o avanço da pesquisa, diversos métodos foram desenvolvidos para a construção de modelos de estimação do risco. Os métodos de *Credit Scoring* geralmente utilizados são baseados em modelos estatísticos e técnicas de reconhecimento de padrão do comportamento do cliente. De acordo com Hand e Henley (1997) e Vojtek e Kočenda (2006) as técnicas mais comumente empregadas são a análise discriminante linear, regressão linear múltipla,

regressão logística, classificador de k vizinhos mais próximos, árvores de classificação e regressão e redes neurais.

Segundo Hand e Henley (1997) em geral não existe um método melhor. Isso depende dos detalhes do problema, dos dados, características e o objetivo da classificação. No entanto, a regressão logística é o método mais favorecido na prática, principalmente devido a quase nenhuma suposição imposta às variáveis (VOJTEK; KOČENDA, 2006). Outra característica que contribui para a utilização da regressão logística é que o resultado pode ser interpretado como uma probabilidade. Esse aspecto é particularmente importante para o modelo CreditRisk+, pois a probabilidade de *default* é um dos principais dados de entrada.

A regressão logística, ou análise *logit*, é uma técnica apropriada para as situações nas quais a variável dependente é binária. Esta assume apenas um entre dois resultados possíveis e, neste caso, avalia-se se um cliente é adimplente ou inadimplente. Dessa forma, a variável resposta é representada com valores 0 ou 1, indicando a ocorrência ou não do evento de interesse (HAIR et al., 2010). O objetivo da regressão logística é gerar um modelo matemático que permite estimar a probabilidade de uma observação pertencer a uma das classes determinadas, sendo explicada por um conjunto de variáveis independentes (BRITO; ASSAF NETO, 2008). As variáveis independentes, por sua vez, representam os fatores que supostamente influenciam no evento de interesse.

Supondo o evento dependente em que a variável resposta é binária e assume valores 0 ou 1 e variáveis independentes x_1, x_2, \dots, x_k , o modelo de regressão logística representa a seguinte relação

$$\ln \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (2.11)$$

em que p é a probabilidade de ocorrer o evento de interesse e β_i os coeficientes estimados. A probabilidade de ocorrência do evento de interesse é obtida pela expressão

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}}. \quad (2.12)$$

2.5 VALOR EM RISCO

O conceito de Valor em Risco (*Value at Risk* - VaR) surgiu na década de 90, motivada pela ocorrência de erros, desastres financeiros e a necessidade de mensurar o risco, tornando-se rapidamente uma métrica de risco em bancos e instituições financeiras (JORION, 2001). Sendo inicialmente utilizado para estimar o risco de mercado, este método de avaliação de risco foi então adaptado para controlar e gerenciar outros tipos de risco, dentre os quais podemos citar o risco de crédito.

Uma definição inicial para o VaR é uma estimativa da perda máxima potencial em um horizonte de tempo, dado um certo nível de confiança (JORION, 2001). Como

ilustração da interpretação de tal medida, considere uma empresa com um VaR diário de R\$10 milhões e 99% de nível de confiança. Isso significa que, no horizonte de um dia, existe apenas 1% de chance de a perda ser superior a R\$10 milhões. Neste caso, pode-se dizer também que com 99% de nível de confiança, a perda não será maior que R\$10 milhões.

Algumas características do VaR a tornaram uma métrica popular na gestão de risco. Dentre elas, podemos citar que o VaR resume em um único número a medida de risco ao qual um ativo ou carteira está exposto, está na mesma unidade de medida analisada e possui interpretação intuitiva, dado que depende apenas do prazo e nível de confiança desejado. Além disso, a métrica é utilizada em regulamentos oficiais na área bancária.

Existem diferentes formas de estimação do VaR e as técnicas variam dependendo da natureza do problema em análise. Estas geralmente envolvem a utilização da distribuição de probabilidade de ganhos ou perdas, ou dados históricos de retornos da carteira sendo analisada. No contexto do presente trabalho, a distribuição de probabilidade de perdas da carteira será gerada pelo modelo CR+. O VaR é, então, um percentil extremo desta distribuição dado um horizonte de tempo, formalmente, sendo definido como

$$VaR_\epsilon[X] = \inf\{c \in \mathbb{R} : P[X \leq c] \geq \epsilon\}, \quad (2.13)$$

onde X representa a perda e ϵ representa um percentil de nível de confiança (em geral entre 95% e 99,9%). Como a distribuição de perda no modelo CR+ é discreta, a estimação do VaR pode ser aproximada e, conforme sugerem Gundlach e Lehrbass (2004), tem-se

$$VaR_\epsilon[X] \approx \bar{\rho}_\epsilon(X) = \inf\{n \in \mathbb{N} : P[X \leq n] \geq \epsilon\}, \quad (2.14)$$

ou ainda por meio de uma interpolação linear

$$VaR_\epsilon[X] \approx \rho_\epsilon(X) = \bar{\rho}_\epsilon(X) - 1 + \frac{\epsilon - P[X \leq \bar{\rho}_\epsilon(X) - 1]}{P[X = \bar{\rho}_\epsilon(X)]}. \quad (2.15)$$

Outra sugestão de interpolação também feita por Gundlach e Lehrbass (2004) considera um percentil de nível de confiança ϵ escolhido e a interpolação é feita entre dois pontos q_1 e q_2 , com $\epsilon \in [q_1, q_2]$

$$VaR_\epsilon = U \left[(v_{A_{q_2}} - 1) + \left(\frac{\epsilon - q_1}{q_2 - q_1} \right) \right] \quad (2.16)$$

onde $v_{A_{q_2}}$ representa o valor de perda discretizada associada à probabilidade acumulada da distribuição de perda agregada até q_2 , e U representa o tamanho da faixa de exposição.

Apesar da popularidade do VaR e sua utilização na gestão de risco na área corporativa, a literatura de risco possui algumas ressalvas quanto ao seu emprego. Uma crítica à esta medida de risco é não fornecer informações sobre a dimensão das perdas em casos extremos onde o prejuízo seja maior que o valor dado pelo VaR e, dessa forma,

consiste apenas em um limite inferior para as perdas (ROCKAFELLAR; URYASEV, 2002).

Esta medida possui também algumas limitações matemáticas ao não possuir propriedades como a diferenciabilidade e convexidade, o que não é ideal em contextos onde o VaR é utilizado como medida de risco na otimização de carteiras, ou problemas de programação matemática em geral. Estas propriedades são essenciais para garantir que os problemas de otimização possuam solução global, afetando contextos onde o VaR é tratado como função objetivo e é necessário minimizar o risco.

O VaR não é considerado uma medida de risco coerente de acordo com a definição de Artzner et al. (1999), pois não é uma medida subaditiva. Não possuir a propriedade de subaditividade significa que o VaR de uma combinação de carteiras pode ser maior que a soma do VaR de cada uma delas, o que significaria que um gestor utilizando esta medida não se beneficia da diversificação das carteiras.

Considerando estas características, uma medida alternativa baseada no VaR foi desenvolvida. Esta é apresentada na próxima seção.

2.6 VALOR EM RISCO CONDICIONAL

O Valor em Risco Condicional (*Conditional Value at Risk* - CVaR), conhecida também como perda máxima esperada ou *Expected Shortfall* (ES) é uma estimativa da perda esperada excedente ao VaR, ou seja, representa o valor médio das perdas no cenário em que estas são maiores que o VaR e quantifica o quão alto é a perda. A título de exemplo, para um VaR com $\epsilon = 0,99$ de nível de confiança, o CVaR é dado pela média de 1% das piores perdas.

Segundo Acerbi e Tasche (2002), o CVaR é uma medida de risco coerente sob a definição em Artzner et al. (1999). Por ser a média das piores perdas, é mais pessimista e maior ou igual ao VaR, podendo também ser considerado um limite superior para as perdas da carteira. Esta métrica de risco possui também as propriedades de ser diferenciável e convexa (ROCKAFELLAR; URYASEV, 2002), o que facilita a sua aplicação em comparação ao VaR.

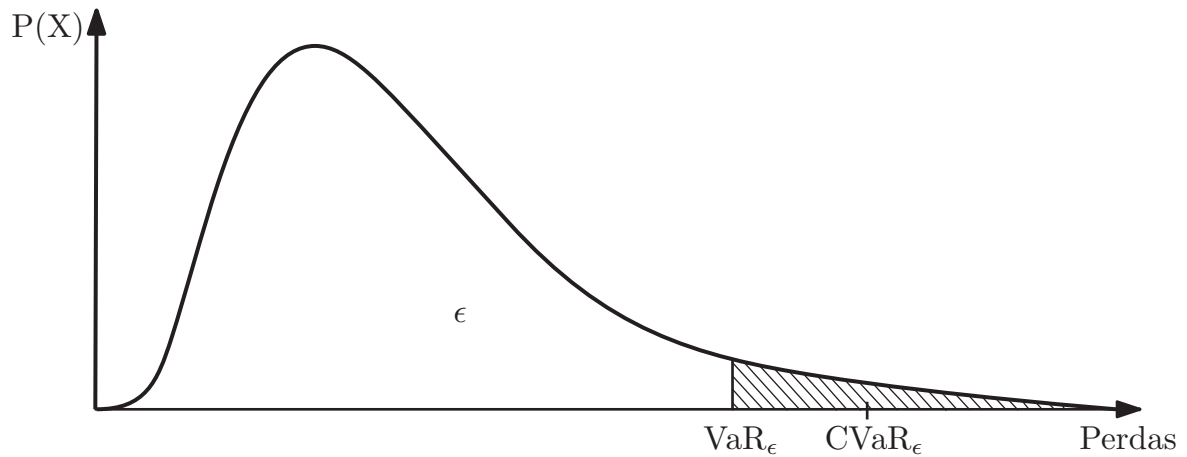
Formalmente, pode ser definida como a esperança das perdas da carteira acima do VaR, sendo escrito como

$$CVaR_{\epsilon}[X] = \mathbb{E}[X : X \geq VaR_{\epsilon}[X]] \quad (2.17)$$

onde X representa a distribuição de perdas da carteira e $VaR_{\epsilon}[X]$ é o VaR com ϵ de nível de confiança.

A Figura 1 apresenta uma ilustração dos valores do VaR e CVaR para uma distribuição de perdas qualquer.

FIGURA 1 – Representação do VaR e CVaR para uma distribuição de perdas



Fonte: A autora (2018)

Dessa forma, o VaR poderia ser utilizado para responder a seguinte pergunta: “Qual a perda máxima potencial ocorrida dado um nível de confiança de $\epsilon\%$?”, enquanto o CVaR responde a “Qual a perda média ocorrida nos $(1 - \epsilon)\%$ piores cenários?”.

2.7 MODELOS DE RISCO DE CARTEIRAS DE CRÉDITO

Diferentes modelos de risco de crédito foram desenvolvidos na década de 90, dentre os quais se destacaram os modelos CreditMetrics, CreditPortfolioView, KMV e CreditRisk+. Apesar de possuírem bases teóricas diferentes, o principal objetivo destes modelos é obter uma distribuição de probabilidades de perda para uma carteira de crédito.

Como discutido anteriormente e descrito a seguir, nem todos os modelos desenvolvidos para estimação de perda agregada são adequados para o mercado financeiro do Brasil. Uma descrição breve de cada modelo, assim como justificativa de não serem compatíveis, são feitas conforme Saunders e Allen (2002), Schechtman et al. (2004), Andrade (2004) e Crouhy, Galai e Mark (2005) .

2.7.1 CreditMetrics

No modelo CreditMetrics, criado pela J.P. Morgan em 1997 (JPM, 1997), o risco é definido pela variação do valor de mercado da carteira, sendo mensurado pela ocorrência de eventos de *default* e principalmente pela mudança na classificação de crédito dos devedores. Dessa forma, a aplicação deste modelo depende de *ratings* de crédito, sendo fornecida por um modelo de classificação interno ou proveniente de uma agência de classificação.

Esta metodologia utiliza o conceito de risco de *spread* e faz uso de matrizes de migração de *rating*. A matriz de migração de *rating* é essencial para a construção deste

modelo, e indica as probabilidades de uma empresa com um determinado *rating* migrar para as outras categorias de classificação ao final do período avaliado. Dessa forma, calcula-se as perdas decorrentes da degradação do *rating* de crédito e dos eventos de *default*.

A distribuição de perdas é então obtida por métodos de simulação como a de Monte Carlo, em que um número alto de possíveis cenários são gerados para compor a distribuição da carteira. Para estimar o valor da carteira nos diversos cenários, o CreditMetrics utiliza as variações de preços das ações de cada devedor da carteira para obter as variações simuladas nas classificações dos devedores.

A aplicação do modelo CreditMetrics no mercado brasileiro é severamente afetada, pois seria necessário um mercado secundário de títulos de crédito, que no Brasil não é muito desenvolvido, para a estimação dos *spreads*. Além disso, o CreditMetrics necessita de um provedor consistente de matrizes de migração de *rating* e a análise seria limitada pelas empresas que são negociadas no mercado de ações.

2.7.2 CreditPortfolioView

O CreditPortfolioView, criado pela McKinsey (WILSON, 1998), é uma metodologia de estimação de risco baseada na relação entre fatores macroeconômicos e a probabilidade de *default* dos devedores. Formalmente, a probabilidade de *default* é descrita em função de diversas variáveis macroeconômicas como o desemprego, o nível da taxa de juros, crescimento da economia e assim por diante. Para geração de dados futuros da probabilidade de *default*, são utilizados modelos de séries temporais auto-regressivos para a previsão das variáveis macroeconômicas. Estes valores futuros das variáveis são então utilizados nos modelos econométricos de *default* de cada segmento para estimação da probabilidade de *default* esperada.

O CreditPortfolioView também utiliza técnicas de simulações para a obtenção da distribuição de perdas da carteira. A aplicação do modelo é feita por meio da simulação de diversos cenários para diferentes valores dos parâmetros. Neste modelo, matrizes de migração de *rating* também são utilizadas. Para cada cenário gerado na simulação, é necessário também avaliar as alterações e migrações de *rating*, sendo que a precificação das carteiras é feita de forma análoga ao modelo CreditMetrics.

Dessa forma, para aplicação deste modelo é necessário uma quantidade considerável de dados sobre as variáveis explicativas, para que seja possível determinar fatores macroeconômicos precisos para a carteira de crédito. No entanto, as mudanças nos regimes macroeconômicos do Brasil tornam a coleta destes dados uma tarefa difícil (ANDRADE, 2004).

2.7.3 KMV

O modelo KMV, desenvolvido pela empresa KMV Corporation (KMV, 1993) supõe que o preço das ações das empresas negociadas no mercado acionário reflete as expectativas do mercado sobre esta empresa e, portanto, pode ser utilizado para precificar o risco de crédito. A metodologia empregada pelo KMV é diferente do CreditMetrics, no sentido que o risco é medido por meio de uma frequência esperada de default (EDF¹), para cada empresa utilizando informações do mercado acionário, em vez de depender de *ratings* de crédito produzidas pelas agências de classificação.

A EDF para cada instituição é uma função da estrutura de capital da empresa, do valor atual de seus ativos e da volatilidade dos retornos dos ativos. O valor dos ativos da empresa é proveniente do valor do mercado de ações.

Dessa forma, o modelo KMV utiliza as informações de preços das ações para a mensuração do risco de inadimplência. Este processo é feito por meio de três etapas principais:

- a) Estimativa do valor de mercado e volatilidade dos ativos da empresa, conforme o mercado de ações.
- b) Cálculo da distância à inadimplência, que é uma medida do índice de risco de inadimplência.
- c) Cálculo da probabilidade de *default* por meio do dimensionamento da distância à inadimplência para uma medida de probabilidades.

A aplicação do modelo KMV é limitada no Brasil por características semelhantes ao do CreditMetrics. Esta metodologia utiliza como dados de entrada parâmetros do mercado acionário, fazendo com que sua aplicação no Brasil seja limitada à grandes empresas com ações negociadas.

2.7.4 CreditRisk+

O modelo CreditRisk+, desenvolvido por Tom Wilde e publicada pela CSFP (1997), é um modelo de origem atuarial que se concentra apenas no evento de *default*. Sua premissa principal é assumir que a distribuição de eventos de *default* e perdas da carteira seguem um processo de Poisson. Ao contrário dos modelos anteriores, o risco de migração de crédito não é explicitamente modelado e, em vez disso, assume-se a existência de fatores de risco que influenciam certos devedores, sendo estes fatores, por sua vez, controlados por distribuições gama independentes.

A base teórica deste modelo possibilita a obtenção de uma solução analítica, assim como o desenvolvimento de uma expressão recursiva para o cálculo da distribuição. Dessa

¹ Expected Default Frequency

forma, não é necessário a utilização de técnicas de simulação. Em comparação aos outros modelos, os dados de entrada exigidos pelo modelo CreditRisk+ são muito menos exigentes e mais acessíveis.

A presente seção se faz breve, pois o próximo capítulo dedica-se à teoria do CreditRisk+, o modelo adotado para a estimação do risco neste trabalho.

3 O MODELO CREDITRISK+

A partir das definições apresentadas nos capítulos anteriores, desde conceitos de teoria de probabilidade até as nomenclaturas de crédito, o presente capítulo apresenta os fundamentos teóricos que sustentam um dos modelos mais utilizados, segundo a literatura, para a determinação da distribuição de perdas, o CreditRisk+.

Como visto anteriormente, o modelo CreditRisk+ tornou-se popular desde sua publicação em 1997 pela Credit Suisse First Boston (CSFB) (CSFP, 1997). A partir da publicação do documento técnico do modelo original, diversas extensões e modificações foram desenvolvidas para viabilizar o CR+ para diferentes cenários e suposições da área de crédito.

Assim, este capítulo apresenta uma base comum à todas as versões do CreditRisk+ e descreve o modelo em sua forma original. Para tanto, os desenvolvimentos apresentados nesta seção seguem os trabalhos e notações de Gundlach e Lehrbass (2004), CSFP (1997) e Avesani et al. (2014).

3.1 CARACTERÍSTICAS E SUPOSIÇÕES

Diferente dos outros modelos de risco de crédito, o CR+ se concentra no evento de *default* para avaliar o risco de uma carteira de investimentos. As características principais do modelo podem ser resumidas, segundo Kluge e Lehrbass (2004), nos seguintes pontos:

- a) Desinteresse pelas causas da inadimplência: A causa do *default* de um devedor não é incluída no modelo, em vez disso a inadimplência é vista como um evento aleatório descrito por uma probabilidade de *default*;
- b) Discretização das perdas: Para que seja possível calcular a perda agregada da carteira de forma eficiente, esta é ajustada à valores inteiros;
- c) Aplicação de funções geradoras: A distribuição de perdas é derivada de funções geradoras de probabilidade;
- d) Aproximação da distribuição de perdas: A função geradora de probabilidades é aproximada por outra distribuição, geralmente de Poisson para a perda de um devedor. Este processo é essencial para obter o modelo em sua forma recursiva;
- e) Correlação entre devedores ocorre por meio de fatores de risco: A correlação entre os devedores surge implicitamente devido aos fatores de risco comuns.

Tais características serão vistas com mais detalhes nas próximas seções, em que o desenvolvimento do modelo CR+ será apresentado.

3.2 COMPONENTES DE ENTRADA DO MODELO CREDITRISK+

Considere uma carteira de crédito de N devedores, com $A = 1, \dots, N$, e a probabilidade de *default* (PD) do devedor A dada por \tilde{p}_A . Seja também V_A a exposição bruta e \tilde{v}_A a exposição líquida do devedor A . A exposição bruta resume todo o valor das operações de crédito, enquanto a exposição líquida considera severidade e possibilidade de recuperação dada inadimplência. Então a perda esperada (PE) para o devedor A é representada por

$$PE_A = \tilde{p}_A \tilde{v}_A. \quad (3.1)$$

A primeira característica que possibilita a eficiência computacional do modelo CR+ é trabalhar com perdas discretas. Neste caso, ocorre a discretização da variável contínua de exposição individual do devedor. A exposição ajustada do devedor A é expressa como o maior valor inteiro v_A e tem-se:

$$v_A = \left\lceil \frac{\tilde{v}_A}{U} \right\rceil \quad (3.2)$$

em que U representa o tamanho da faixa, e é denominado também como a unidade de perda. No modelo CR+ é determinado por

$$U = \frac{\max[\tilde{v}_A]}{F} \quad (3.3)$$

em que F é o número de faixas adotado.

Este processo de discretização limita a quantidade de valores possíveis de perda, o que reduz o tempo de estimação da distribuição de perdas. Entretanto, o modelo é sensível à quantidade de faixas, onde um número muito grande (e no caso extremo ser a maior \tilde{v}_A) o tornaria inviável e um número muito pequeno prejudicaria a estimação da distribuição final.

Cabe notar que a separação em faixas introduz uma aproximação no início do cálculo. No entanto, desde que o número de exposições seja grande e o tamanho das faixas seja pequeno em comparação com a exposição média da carteira, tal aproximação é insignificante (CSFP, 1997).

Uma solução possível para compensar tal erro no processo de arredondamento é ajustar a PD para preservar o valor da PE_A de forma que

$$p_A = \frac{PE_A}{v_A U}, \quad (3.4)$$

onde tem-se a relação $\tilde{p}_A \tilde{v}_A = U p_A v_A$.

Uma vez que as exposições individuais tenham sido ajustadas e arredondadas, conforme a equação (3.2), elas precisam ser agregadas por faixas de exposição. Sendo F o

número total de faixas e f representando cada faixa de exposição, o valor da exposição na faixa f é $v_f = f$. O devedor A é atribuído à faixa f se $v_A = v_f = f$. Dessa forma, a probabilidade de *default* na faixa f é dada por

$$p_f = \sum_{A, v_A=f} p_A. \quad (3.5)$$

Como os devedores foram agregados por faixa de exposição, p_A descreve uma expectativa de inadimplência durante um certo período, em uma determinada condição econômica, para devedores com mesmo perfil de risco. Com o propósito de considerar o efeito de devedores pertencentes a uma indústria específica ou afetados por diferentes cenários econômicos, ou ainda serem de países diferentes, o devedor A tem a p_A ajustada por um fator escalar que representa o risco específico do cenário em questão. Este processo no modelo CR+ é feito por meio da inclusão de setores, sendo assim, inserindo o risco proveniente das características diferentes dos devedores.

Para tanto, considere K diferentes e independentes setores s_1, \dots, s_K com seus respectivos fatores aleatórios escalares S_k , $k = 1, \dots, K$, que são chamados de fatores de risco. Cada devedor A pode estar ativo em mais de um setor e inclui-se um peso w_{Ak} para representar essa participação no setor k . Tais pesos indicam como a probabilidade de *default* do devedor A é influenciada pelo k -ésimo setor. Estes pesos também precisam satisfazer

$$0 \leq w_{Ak} \leq 1 \quad \text{e} \quad \sum_{k=1}^K w_{Ak} \leq 1. \quad (3.6)$$

No qual as seguintes expressões representam os riscos envolvidos

- $\sum_{k=1}^K w_{Ak}$ é o risco sistemático de *default* do devedor A ;
- $w_{A0} = 1 - \sum_{k=1}^K w_{Ak}$ representa o risco idiossincrático de *default* do devedor A .

O conceito de risco idiossincrático pode ser definido como o risco específico e inerente a um devedor, a parte da inadimplência que não pode ser explicada pela inclusão de setores. Note que se não houver risco idiossincrático, sendo $w_{A0} = 0$, existe apenas o risco sistemático e $\sum_{k=1}^K w_{Ak} = 1$.

Para a inclusão de setores no modelo, supõe-se que a relação entre os fatores de risco e a probabilidade de *default* é linear, desse modo, p_A é ajustada novamente para considerar a influência dos setores.

Por hipótese, os fatores de risco S_k são assumidos como variáveis aleatórias independentes com distribuição gama, $S_k \sim \text{Gama}(\alpha_k, \beta_k)$. Condicionamente aos $S =$

(S_0, S_1, \dots, S_k) , com $S_0 = 1$, a PD é ajustada novamente por

$$p_A^S = p_A \sum_{k=0}^K w_{Ak} \frac{S_k}{\mathbb{E}[S_k]}, \quad (3.7)$$

com

$$\mathbb{E}[p_A^S] = p_A \quad \text{e} \quad \text{var}[p_A^S] = p_A^2 \sum_{k=1}^K \frac{w_{Ak}^2}{\mathbb{E}[S_k]^2} \text{var}[S_k]. \quad (3.8)$$

Com a introdução de setores, outra suposição do modelo CR+ é assumir que, condicionalmente a S , os eventos de *default* dos devedores são independentes entre si. Dessa forma, as correlações entre os devedores não são diretas e explícitas, mas surgem implicitamente devido aos fatores de riscos comuns que modificam a probabilidade de *default*.

Conhecidos os principais componentes de entrada do modelo CR+, a próxima seção apresenta como a distribuição de perda da carteira de crédito é obtida.

3.3 DISTRIBUIÇÃO DE PERDA

Para um devedor A a perda agregada pode ser representada por uma variável aleatória $X_A = \mathbb{I}_A v_A$, onde

$$\mathbb{I}_A = \begin{cases} 1, & \text{se } A \text{ entra em } \textit{default} \\ 0, & \text{caso contrário} \end{cases}. \quad (3.9)$$

O evento de *default* é representado por uma variável aleatória Bernoulli, e assume apenas dois estados: inadimplência ou adimplência. A distribuição de probabilidade de X_A é dada por $P(X_A = v_A) = p_A$ e $P(X_A = 0) = 1 - p_A$. Assim, a probabilidade de perder o valor em exposição é dada pela PD e a probabilidade de não haver perdas é 1-PD.

O objetivo do modelo CR+ é determinar a distribuição de probabilidades da perda da carteira de forma computacionalmente eficiente a fim de estimar medidas de risco como o VaR. Para tanto, a distribuição de perdas é derivada da função geradora de probabilidade, onde X_A assume apenas valores inteiros não negativos.

Dessa forma, a distribuição de perda de um devedor A é escrita por meio da FGP em (2.2) como

$$G_A(z) = (1 - p_A)z^0 + p_A z^{v_A} = 1 + p_A(z^{v_A} - 1). \quad (3.10)$$

Para a distribuição de perda de um devedor A , condicional aos setores S , a FGP é dada por

$$G_A(z|S) = (1 - p_A^S)z^0 + p_A^S z^{v_A} = 1 + p_A^S(z^{v_A} - 1). \quad (3.11)$$

Tal expressão pode ser reescrita também como

$$G_A(z|S) = e^{\ln(1+p_A^S(z^{v_A}-1))}. \quad (3.12)$$

Para se obter a distribuição de perdas da carteira de forma eficiente, a FGP (3.11) é aproximada por uma distribuição de Poisson no lugar da distribuição de Bernoulli. Para que esta aproximação seja possível, é essencial supor que as probabilidades individuais de *default* são suficientemente pequenas.

Como a distribuição dos eventos de *default* é dada por uma variável aleatória Bernoulli, e assumindo que p_A^S é pequena, então $p_A^S(z^{v_A} - 1)$ também será suficientemente pequeno desde que seja válido $|z| \leq 1$. Definindo $w = p_A^S(z^{v_A} - 1)$, e utilizando a expansão de Taylor da função $\ln(1 + w)$ na vizinhança de $w = 0$ tem-se

$$\ln(1 + w) = w - \frac{w^2}{2} + \frac{w^3}{3} \dots \quad (3.13)$$

Considerando w suficientemente pequeno e desprezando os termos de ordem maiores ou iguais a 2, é válida a aproximação

$$\ln(1 + w) \approx w. \quad (3.14)$$

Portanto, de (3.12) e (3.14), e escrevendo w em sua notação original, tem-se

$$G_A(z|S) \approx e^{p_A^S(z^{v_A}-1)}. \quad (3.15)$$

Ainda, pela expansão de Taylor da expressão $e^{p_A^S(z^{v_A}-1)}$ na vizinhança de $z = 0$ tem-se

$$e^{p_A^S(z^{v_A}-1)} = e^{-p_A^S} + e^{-p_A^S} p_A^S \frac{z}{1!} + e^{-p_A^S} (p_A^S)^2 \frac{z^2}{2!} + e^{-p_A^S} (p_A^S)^3 \frac{z^3}{3!} + \dots \quad (3.16)$$

Conclui-se então que

$$G_A(z|S) \approx e^{p_A^S(z^{v_A}-1)} = e^{-p_A^S} \sum_{x=0}^{\infty} \frac{(p_A^S)^x}{x!} z^x, \quad (3.17)$$

onde a expressão (3.15) é a FGP de uma distribuição de Poisson com taxa p_A^S . Portanto, considera-se como válida a aproximação da distribuição de Bernoulli pela distribuição de Poisson, desde que p_A^S seja pequena. Como os eventos de *default* dos devedores são independentes condicionalmente aos setores S , tem-se que a FGP de X é dada por

$$G_X(z|S) = \prod_A G_A(z|S) = \exp \left(\sum_A p_A^S(z^{v_A} - 1) \right). \quad (3.18)$$

Assumindo sem perda de generalidade que as distribuições dos fatores S_i sejam tais que $\mathbb{E}[S_i] = 1$ (o que garante que $\mathbb{E}[p_S^A] = p_A$) e da expressão da p_A em (3.7) tem-se

$$G_X(z|S) = \exp \left(\sum_A \sum_{k=0}^K p_A w_{Ak} S_k(z^{v_A} - 1) \right). \quad (3.19)$$

Definindo μ_k e o polinômio setorial \mathcal{P}_k como

$$\mu_k = \sum_A w_{Ak} p_A \quad (3.20)$$

$$\mathcal{P}_k(z) = \frac{1}{\mu_k} \sum_A w_{Ak} p_A z^{v_A} \quad (3.21)$$

reescreve-se a expressão (3.19) como

$$\begin{aligned} G_X(z|S) &= \exp \left(\sum_{k=0}^N S_k \left(\sum_A p_A w_{Ak} (z^{v_A} - 1) \right) \right) \\ &= \exp \left(\sum_{k=0}^N S_k \mu_k (\mathcal{P}_k(z) - 1) \right). \end{aligned} \quad (3.22)$$

Representando a função densidade de probabilidade dos fatores de risco S_k por f_S , também é possível determinar a função geradora de X como uma distribuição composta

$$G_X(z) = \int G_X(z|S=s) f_S(s) ds. \quad (3.23)$$

Neste modelo, a distribuição gama é utilizada para os fatores de risco S_k . Como foi assumido que $\mathbb{E}[S_i] = 1$, pode-se estimar os parâmetros da distribuição de S_k . Sendo $S_k \sim \text{Gama}(\alpha_k, \beta_k)$, conclui-se que

$$\begin{aligned} \mathbb{E}[S_k] &= \alpha_k \beta_k = 1 \\ \text{var}[S_k] &= \alpha_k \beta_k^2 = \sigma_k^2, \end{aligned} \quad (3.24)$$

logo,

$$\begin{aligned} \alpha_k &= \frac{1}{\sigma_k^2} \\ \beta_k &= \sigma_k^2. \end{aligned}$$

Como $\mathbb{E}[S_k] = \mu_k$, pode-se escrever também

$$\begin{aligned} \alpha_k &= \frac{\mu_k^2}{\sigma_k^2} \\ \beta_k &= \frac{\sigma_k^2}{\mu_k}. \end{aligned} \quad (3.25)$$

Conhecendo a distribuição de f_S dos fatores de risco S_k e seus parâmetros, é possível obter uma fórmula analítica fechada quando f_S é assumida como uma distribuição gama. Neste caso, de (3.23)

$$G_X(z) = \int G_X(z|S=s) f_S^{(\alpha, \beta)}(s) ds. \quad (3.26)$$

Portanto, de (2.8) e (3.22), assumindo a independência entre os K setores e utilizando integração múltipla, com $s_0 \equiv 1$ e parâmetros $\alpha = (\alpha_1, \dots, \alpha_K)$ e $\beta = (\beta_1, \dots, \beta_K)$, escreve-se

$$G_X(z) = \int \exp \left(\sum_{k=0}^N S_k \mu_k (\mathcal{P}_k(z) - 1) \right) \prod_{l=1}^K \frac{s_l^{\alpha_l - 1} e^{-s_l / \beta_l}}{\beta_l^{\alpha_l} \Gamma(\alpha_l)} ds_l. \quad (3.27)$$

Uma vez que $\beta_k = \frac{1}{\alpha_k}$

$$\begin{aligned} G_X(z) &= \int \exp\left(\sum_{k=0}^N S_k \mu_k (\mathcal{P}_k(z) - 1)\right) \prod_{l=1}^K \frac{s_l^{\alpha_l-1} e^{-s_l \alpha_l} \alpha_l^{\alpha_l}}{\Gamma(\alpha_l)} ds_l \\ &= e^{\mu_0(\mathcal{P}_0(z)-1)} \prod_{k=1}^K \frac{\alpha_k^{\alpha_k}}{\Gamma(\alpha_k)} \int_0^\infty e^{s_k(\mu_k(\mathcal{P}_k(z)-1)-\alpha_k)} s_k^{\alpha_k-1} ds_k. \end{aligned}$$

Como

$$\frac{\Gamma(b)}{a^b} = \int_0^\infty t^{b-1} e^{-at} dt,$$

tem-se

$$\begin{aligned} G_X(z) &= e^{\mu_0(\mathcal{P}_0(z)-1)} \prod_{k=1}^K \frac{\alpha_k^{\alpha_k}}{\Gamma(\alpha_k)} \frac{\Gamma(\alpha_k)}{[\alpha_k - \mu_k(\mathcal{P}_k(z) - 1)]^{\alpha_k}} \\ &= e^{\mu_0(\mathcal{P}_0(z)-1)} \prod_{k=1}^K \left(\frac{\alpha_k}{\alpha_k - \mu_k \mathcal{P}_k(z) + \mu_k} \right)^{\alpha_k} \\ &= e^{\mu_0(\mathcal{P}_0(z)-1)} \prod_{k=1}^K \left(\frac{\frac{\alpha_k}{\mu_k + \alpha_k}}{\frac{\alpha_k}{\mu_k + \alpha_k} - \frac{\mu_k \mathcal{P}_k(z)}{\mu_k + \alpha_k} + \frac{\mu_k}{\mu_k + \alpha_k}} \right)^{\alpha_k}. \end{aligned}$$

Definindo

$$\delta_k = \frac{\mu_k}{\mu_k + \alpha_k} \quad (3.28)$$

e notando que

$$\frac{\alpha_k}{\mu_k + \alpha_k} = 1 - \frac{\mu_k}{\mu_k + \alpha_k},$$

pode-se escrever a função geradora como

$$G_X(z) = e^{\mu_0(\mathcal{P}_0(z)-1)} \prod_{k=1}^K \left(\frac{1 - \delta_k}{1 - \delta_k \mathcal{P}_k(z)} \right)^{\alpha_k}. \quad (3.29)$$

Dessa forma, a distribuição de perdas para cada setor é obtida como uma mistura contínua da distribuição de Poisson como composta da Gama. Isto é, a distribuição de perdas do setor é uma Poisson com fator de risco com distribuição Gama, o que resulta numa distribuição binomial negativa para cada setor. A distribuição agregada de perdas da carteira, por sua vez, não é binomial negativa, mas é construída como uma soma de distribuições independentes desse tipo.

Cabe notar que diferentes funções podem ser utilizadas para estimar a distribuição de perdas da carteira. Enquanto a publicação original utiliza funções geradoras de probabilidade, Giese (2003) aborda o problema com a utilização de funções geradoras de momentos, o que oferece a possibilidade de modelagem de correlações entre os setores. O método de Aproximação Ponto de Sela, proposta por Gordy (2002), é baseada na função geradora de cumulantes e Reiß (2004) apresenta uma forma de estimação utilizando a função característica.

3.4 FORMA RECURSIVA

Para obter a distribuição de probabilidades seria necessário, conforme (2.2), calcular as n -ésimas derivadas da expressão G_X para obter $P(X = n)$ para cada valor de perda n . Com o propósito de estimar a distribuição de perdas de uma carteira pela FGP de forma computacionalmente eficiente, foi sugerido no documento original do modelo CR+ a utilização de um processo recursivo conhecido como recursividade de Panjer (PANJER, 1981). Essa recursividade no modelo é baseada na representação da derivada de $\ln G_X(z)$ como a razão entre dois polinômios $A(z)$ e $B(z)$, desenvolvido como a seguir.

Retomando de (3.29), pode-se escrever

$$\ln G_X(z) = \mu_0(\mathcal{P}_0(z) - 1) + \sum_{k=1}^K \alpha_k \ln \left(\frac{1 - \delta_k}{1 - \delta_k \mathcal{P}_k(z)} \right). \quad (3.30)$$

Desta forma, tem-se então

$$\begin{aligned} (\ln G_X)'(z) &= \mu_0 \mathcal{P}_0'(z) + \sum_{k=1}^K \alpha_k \frac{1 - \delta_k \mathcal{P}_k(z)}{1 - \delta_k} \frac{1 - \delta_k}{(1 - \delta_k \mathcal{P}_k(z))^2} \delta_k \mathcal{P}_k'(z) \\ &= \mu_0 \mathcal{P}_0'(z) + \sum_{k=1}^K \alpha_k \frac{\delta_k \mathcal{P}_k'(z)}{1 - \delta_k \mathcal{P}_k(z)} \end{aligned} \quad (3.31)$$

Definindo $\eta_0 = \mu_0$ e $\eta_k = \alpha_k \delta_k$,

$$\eta_k = \alpha_k \delta_k = \alpha_k \frac{\mu_k}{\mu_k + \alpha_k} = \frac{\mu_k}{1 + \beta \mu_k} \quad (3.32)$$

visto que $\alpha_k \beta_k = 1$ e $\alpha_k = \frac{1}{\beta}$.

Note também que

$$\mathcal{P}_0'(z) = \frac{\mathcal{P}_0'(z)}{1 - \delta_0 \mathcal{P}_0(z)} \Leftrightarrow \delta_0 = 0. \quad (3.33)$$

Logo reescreve-se a expressão (3.31) como

$$\begin{aligned} (\ln G_X)'(z) &= \eta_0 \frac{\mathcal{P}_0'(z)}{1 - \delta_0 \mathcal{P}_0(z)} + \sum_{k=1}^K \eta_k \frac{\mathcal{P}_k'(z)}{1 - \delta_k \mathcal{P}_k(z)} \\ &= \sum_{k=0}^K \eta_k \frac{\mathcal{P}_k'(z)}{1 - \delta_k \mathcal{P}_k(z)} \end{aligned} \quad (3.34)$$

A título de exemplo, considere a expressão anterior para $K = 2$

$$\begin{aligned} (\ln G_X)'(z) &= \sum_{k=0}^2 \eta_k \frac{\mathcal{P}_k'(z)}{1 - \delta_k \mathcal{P}_k(z)} \\ &= \eta_0 \frac{\mathcal{P}_0'(z)}{1} + \eta_1 \frac{\mathcal{P}_1'(z)}{1 - \delta_1 \mathcal{P}_1(z)} + \eta_2 \frac{\mathcal{P}_2'(z)}{1 - \delta_2 \mathcal{P}_2(z)}. \end{aligned} \quad (3.35)$$

Assim

$$\begin{aligned}
 (\ln G_X)'(z) &= \frac{\eta_0 \mathcal{P}'_0(z)(1 - \delta_1 \mathcal{P}_1(z))(1 - \delta_2 \mathcal{P}_2(z))}{(1 - \delta_1 \mathcal{P}_1(z))(1 - \delta_2 \mathcal{P}_2(z))} \\
 &+ \frac{\eta_1 \mathcal{P}'_1(z)(1 - \delta_2 \mathcal{P}_2(z))}{(1 - \delta_1 \mathcal{P}_1(z))(1 - \delta_2 \mathcal{P}_2(z))} \\
 &+ \frac{\eta_2 \mathcal{P}'_2(z)(1 - \delta_1 \mathcal{P}_1(z))}{(1 - \delta_1 \mathcal{P}_1(z))(1 - \delta_2 \mathcal{P}_2(z))}.
 \end{aligned}$$

Considerando este desenvolvimento para um K qualquer, a ideia principal da recursividade de Panjer é escrever a expressão (3.34) como

$$(\ln G_X)'(z) = \frac{\sum_{k=0}^K \eta_k \mathcal{P}'_k(z) \prod_{j=0, j \neq k}^K (1 - \delta_j \mathcal{P}_j(z))}{\prod_{l=0}^K (1 - \delta_l \mathcal{P}_l(z))} = \frac{A(z)}{B(z)} \quad (3.36)$$

isto é,

$$(\ln G_X)'(z) = \frac{A(z)}{B(z)} = \frac{a_0 + a_1 z + \dots + a_{d_A} z^{d_A}}{b_0 + b_1 z + \dots + b_{d_B} z^{d_B}} = \frac{\sum_{j=0}^{d_A} a_j z^j}{\sum_{j=0}^{d_B} b_j z^j}. \quad (3.37)$$

O grau desses polinômios, representado pela notação \deg , podem ser determinados por

$$\begin{aligned}
 d_A &= \deg(A) = \sum_{j=0}^K \deg(\mathcal{P}_j - 1) \\
 d_B &= \deg(B) = \sum_{j=1}^K \deg(\mathcal{P}_j).
 \end{aligned} \quad (3.38)$$

Note que, pela definição de polinômio setorial em (3.21), o grau de \mathcal{P}_j é dado pelo maior valor de perda de um devedor qualquer do setor j . Ou seja é a v_A do devedor com a maior exposição pertencente ao setor j . Além disso, como o polinômio B não possui o termo $\mathcal{P}_0(z)$, pois $\delta_0 = 0$ de (3.33), os índices iniciais dos somatórios em (3.38) são diferentes para cada polinômio.

Os coeficientes a_n e b_n podem ser obtidos dos polinômios pela diferenciação

$$\begin{aligned}
 a_n &= \frac{1}{n!} A^{(n)}(0) = \frac{1}{n!} \left. \frac{d^n A}{dz^n} \right|_{z=0} \\
 b_n &= \frac{1}{n!} B^{(n)}(0) = \frac{1}{n!} \left. \frac{d^n B}{dz^n} \right|_{z=0}
 \end{aligned} \quad (3.39)$$

e utilizando a regra de Leibniz, conforme em (KLUGE; LEHRBASS, 2004)

$$\frac{1}{n!} \left(\prod_{k=0}^K h_k \right)^{(n)} = \sum_{n_1 + \dots + n_K = n} \prod_{k=0}^K \frac{h_k^{(n_k)}}{n_k!} \quad (3.40)$$

para qualquer função diferenciável h_k .

Este processo é semelhante ao empregado para a obtenção da distribuição de probabilidades, o qual será apresentado no decorrer deste capítulo. Para a distribuição de probabilidades de perda da carteira, de (2.3), utiliza-se

$$P(X = n) = g_n = \frac{1}{n!} G_X^{(n)}(0). \quad (3.41)$$

Neste momento, é necessário analisar o polinômio setorial \mathcal{P}_k e suas derivadas para estimar os coeficientes a_n e b_n . Apenas recordando (3.21),

$$\mathcal{P}_k(z) = \frac{1}{\mu_k} \sum_A w_{Ak} p_A z^{v_A}$$

tem-se

$$\mathcal{P}_k(0) = 0. \quad (3.42)$$

Enquanto a r -ésima derivada do polinômio setorial é dada por

$$\mathcal{P}_k^{(r)}(z) = \frac{1}{\mu_k} \sum_{A: v_A > r} w_{Ak} p_A v_A (v_A - 1) \cdots (v_A - r + 1) z^{v_A - r}, \quad (3.43)$$

$$\mathcal{P}_k^{(r)}(0) = \frac{1}{\mu_k} \sum_{A: v_A = r} w_{Ak} p_A v_A r!. \quad (3.44)$$

Definindo

$$c_{r,k} = \sum_{A: v_A = r} w_{Ak} p_A \quad (3.45)$$

tem-se

$$\mathcal{P}_k^{(r)}(0) = \frac{1}{\mu_k} c_{r,k} r!. \quad (3.46)$$

Para a estimação dos coeficientes dos polinômios $A(z)$ e $B(z)$ define-se também

$$\delta_{ij} = \begin{cases} 1, & \text{caso } i = j \\ 0, & \text{caso contrário} \end{cases} \quad (3.47)$$

$$\mathbb{I}(n) = \begin{cases} 1, & \text{caso } n = v_A \text{ para algum } A \\ 0, & \text{caso contrário} \end{cases} \quad (3.48)$$

em que δ_{ij} é a função delta de Kronecker e $\mathbb{I}(n)$ é a função indicadora (ou função característica) do conjunto de perdas da carteira.

Portanto, os coeficientes b_n são obtidos por

$$\begin{aligned}
b_n &= \frac{1}{n!} \left(\prod_{k=1}^K (1 - \delta_k \mathcal{P}_k) \right)^{(n)} (0) \\
&= \sum_{n_1 + \dots + n_K = n} \prod_{k=1}^K \frac{(1 - \delta_k \mathcal{P}_k)^{(n_k)}}{n_k!} (0) \\
&= \sum_{n_1 + \dots + n_K = n} \prod_{k=1}^K \left(\delta_{0n_k} - \mathbb{I}(n_k) \delta_k \frac{\mathcal{P}_k^{(n_k)}}{n_k!} \right) \\
&= \sum_{n_1 + \dots + n_K = n} \prod_{k=1}^K \left(\delta_{0n_k} - \mathbb{I}(n_k) \delta_k \frac{c_{n_k, k}}{\mu_k} \right). \tag{3.49}
\end{aligned}$$

De forma análoga, a expressão para os coeficientes a_n é dada por

$$a_n = \sum_{k=0}^K \sum_{n_0 + \dots + n_K = n} \eta_k \frac{c_{n_{k+1}, k}}{\mu_k} (n_k + 1) \mathbb{I}(n_k + 1) \prod_{j=0, j \neq k}^K \left(\delta_{0n_j} - \mathbb{I}(n_j) \delta_j \frac{c_{n_j, j}}{\mu_j} \right). \tag{3.50}$$

Determinados os coeficientes a_n e b_n , o processo recursivo para determinação da distribuição de perdas é apresentado nas expressões seguintes. Primeiro, observe as relações da derivada da função \ln

$$(\ln G_X)'(z) = \frac{1}{G_X(z)} G_X'(z) \Leftrightarrow G_X'(z) = G_X(z) (\ln G_X)'(z). \tag{3.51}$$

Como a ideia principal da recursividade de Panjer é escrever $(\ln G_X)'$ como quociente de polinômios

$$G_X'(z) = G_X(z) (\ln G_X)'(z) = G_X(z) \frac{A(z)}{B(z)}, \tag{3.52}$$

$$\begin{aligned}
G_X''(z) &= \frac{1}{B(z)} [G_X(z) A'(z) + G_X'(z) A(z)] - G_X(z) A(z) \frac{B'(z)}{B^2(z)} \\
&= \frac{1}{B(z)} [G_X(z) A'(z) + G_X'(z) A(z)] - G_X'(z) \frac{B'(z)}{B(z)} \\
&= \frac{1}{B(z)} [G_X(z) A'(z) + G_X'(z) [A(z) - B'(z)]] . \tag{3.53}
\end{aligned}$$

Generalizando para a derivada de ordem n , segundo (KLUGE; LEHRBASS, 2004), tem-se

$$\begin{aligned}
G_X^{(n)}(z) &= \frac{1}{B(z)} \left[G_X(z) A^{(n-1)}(z) \right. \\
&\quad \left. + \sum_{j=1}^{n-1} G_X^{(j)} \left[\binom{n-1}{j} A^{(n-1-j)}(z) - \binom{n-1}{n-j} B^{(n-j)}(z) \right] \right]. \tag{3.54}
\end{aligned}$$

Os dois primeiros termos da distribuição de perdas são, por meio de (3.30) e (3.52), dados por

$$\begin{aligned} g_0 &= G_X(0) = e^{\ln G_X(0)} = \exp \left(\mu_0(\mathcal{P}_0(0) - 1) + \sum_{k=1}^K \alpha_k \ln \left(\frac{1 - \delta_k}{1 - \delta_k \mathcal{P}_k(0)} \right) \right) \\ &= e^{-\mu_0} \prod_{k=1}^K (1 - \delta_k)^{\alpha_k} \end{aligned} \quad (3.55)$$

$$\begin{aligned} g_1 &= \frac{G'_X(0)}{1!} = G_X(0) \frac{A(0)}{B(0)} \\ &= g_0 \frac{a_0}{b_0}. \end{aligned} \quad (3.56)$$

De (3.39) pode-se escrever as seguintes relações

$$\begin{aligned} a_n &= \frac{1}{n!} A^{(n)}(0) \Leftrightarrow A^{(n)}(0) = n! a_n \\ b_n &= \frac{1}{n!} B^{(n)}(0) \Leftrightarrow B^{(n)}(0) = n! b_n \end{aligned} \quad (3.57)$$

e é possível determinar que

$$\begin{aligned} A^{(n-1)}(0) &= (n-1)! a_{n-1} \\ A^{(n-1-j)}(0) &= (n-1-j)! a_{n-1-j} \\ B^{(n-j)}(0) &= (n-j)! b_{n-j}. \end{aligned} \quad (3.58)$$

Dessa forma, a expressão de recursividade é determinada utilizando as relações anteriores, (3.58), a expressão (3.54) e fazendo

$$\begin{aligned} g_n &= \frac{1}{n!} G_X^{(n)}(0) \\ &= \frac{1}{n! b_0} \left[g_0 (n-1)! a_{n-1} \right. \\ &\quad \left. + \sum_{j=1}^{n-1} j! g_j \left[\binom{n-1}{j} (n-1-j)! a_{n-1-j} - \binom{n-1}{n-j} (n-j)! b_{n-j} \right] \right]. \end{aligned} \quad (3.59)$$

Note que em (3.54) são calculadas as derivadas de G_X de ordem 1 até $n-1$. Ao mesmo tempo, as n -ésimas derivadas de $A(z)$ e $B(z)$ não podem ser de ordem superior a, respectivamente, d_A e d_B . Pois neste caso, ocorreria a derivada de valores constantes, o que tornaria o valor da expressão zero. Portanto, é necessário que $n-1 = d_A \Leftrightarrow n - d_A = 1$ e se escreve o somatório $\sum_{j=1}^{n-1}$ como $\sum_{j=\max(1, n-d_A)}^{n-1}$. O mesmo processo deve ser feito para o polinômio B , como supõe-se que $\delta_0 = 0$ em (3.33) e o polinômio em (3.36) não possui o termo $\mathcal{P}_0(z)$, faz-se $n-1 = d_B + 1 \Leftrightarrow n-1-d_B = 1$, reescrevendo o somatório $\sum_{j=1}^{n-1}$

como $\sum_{j=\max(1, n-1-d_B)}^{n-1}$. Assim,

$$g_n = \frac{1}{n! b_0} \left(g_0 (n-1)! a_{n-1} + \sum_{\max(1, n-d_A)}^{n-1} j! g_j \frac{(n-1)!}{j! (n-1-j)!} (n-1-j)! a_{n-1-j} - \sum_{\max(1, n-1-d_B)}^{n-1} j! g_j \frac{(n-1)!}{(n-j)! (j-1)!} (n-j)! b_{n-j} \right). \quad (3.60)$$

Fazendo $k = n - j$ e $n! = n(n-1)!$

$$g_n = \frac{1}{n b_0} \left(g_0 a_{n-1} + \sum_{k=1}^{\min(d_A, n-1)} g_{n-k} a_{k-1} - \sum_{k=1}^{\min(d_B, n-1)} g_{n-k} (n-k) b_k \right) \quad (3.61)$$

e $n = n + 1$ e $k = j + 1$

$$g_n = \frac{1}{(n+1) b_0} \left(g_0 a_n + \sum_{j=1}^{\min(d_A, n)} g_{n-j} a_j - \sum_{k=1}^{\min(d_B, n)} (n-j) g_{n-j} b_{j+1} \right). \quad (3.62)$$

Com a expressão de recursividade dada por

$$g_{n+1} = \frac{1}{b_0(n+1)} \left(\sum_{j=0}^{\min(d_A, n)} a_j g_{n-j} - \sum_{j=0}^{\min(d_B, n)-1} (n-j) b_{j+1} g_{n-j} \right). \quad (3.63)$$

A expressão (3.63) define a distribuição de probabilidade da perda da carteira. Para a determinação da distribuição de perdas, o algoritmo é implementado de forma a calcular a distribuição até que a soma acumulada de g se aproxime de 1, isto é, até $G(j) = g_0 + g_1 + \dots + g_j$ atingir um percentil admissível (entre 95% e 99,9%) ou até um número máximo de iterações ser alcançado.

3.5 EXEMPLOS DA APLICAÇÃO DO MODELO CREDITRISK+

Esta seção apresenta dois exemplos da aplicação da teoria do modelo CR+ em carteiras de crédito hipotéticas. O objetivo destas ilustrações é elucidar os conceitos apresentados nas seções anteriores, sejam dos dados de entrada, os ajustes realizados nos dados, até as expressões utilizadas para a computação da distribuição de perda agregada.

O primeiro exemplo apresentado visa o entendimento da estimação da distribuição agregada de perdas por meio das funções geradoras de probabilidade. Enquanto o segundo exemplo ilustra a aplicação do modelo CR+ de forma manual e detalhada.

3.5.1 Distribuição de Bernoulli e Função Geradora de Probabilidade

Considere uma carteira de crédito hipotética com dois devedores. Assumindo que eles estão em um único setor e que existe apenas o risco sistemático, tem-se $K = 1$ e

$w_{A0} = 0$, $w_{Ak} = 1$ para $A = 1, 2$. Adotando o número de faixas $F = 2$, a Tabela 1 apresenta os dados de entrada do exemplo¹.

TABELA 1 – Exemplo 1 - Dados da carteira

A	\tilde{p}_A	\tilde{v}_A	p_A	v_A
1	0,20	150	0,15	1
2	0,10	400	0,1	2

A unidade de perda e tamanho da faixa U desta carteira é determinada por meio de (3.3)

$$U = \frac{\max [\tilde{v}_A]}{F} = \frac{400}{2} = 200. \quad (3.64)$$

A exposição discretizada v_A e a probabilidade ajustada p_A são determinadas por (3.2) e (3.4). O exemplo a seguir mostra o cálculo para o devedor 2:

$$\begin{aligned} v_2 &= \left\lceil \frac{\tilde{v}_2}{U} \right\rceil = \left\lceil \frac{400}{200} \right\rceil = 2 \\ p_2 &= \frac{PE_2}{v_2 U} = \frac{\tilde{p}_2 \tilde{v}_2}{v_2 U} = \frac{0,10 \cdot 400}{2 \cdot 200} = 0,10. \end{aligned} \quad (3.65)$$

Ao final de um período, cada devedor A pode estar em uma de duas situações possíveis: em adimplência ou inadimplência. Neste caso, representando a perda monetária por meio da variável aleatória X_A , e considerando que a taxa de recuperação dado *default* é zero, a distribuição de probabilidade de perdas de um devedor é dada por

$$P(X_A = v_A) = p_A \quad P(X = 0) = 1 - p_A. \quad (3.66)$$

Recordando de (2.2) a função geradora de probabilidades de um devedor A é dada por

$$G_A(z) = (1 - p_A)z^0 + p_A z^{v_A}. \quad (3.67)$$

Supondo que os eventos de *default* dos devedores são independentes, a função geradora de perda agregada da carteira é dada por

$$\begin{aligned} G_X(z) &= \prod_{A=1}^2 G_A(z) \\ &= (0,85z^0 + 0,15z^1)(0,90z^0 + 0,10z^2) \\ &= 0,765z^0 + 0,135z^1 + 0,085z^2 + 0,015z^3. \end{aligned} \quad (3.68)$$

Segundo a definição de função geradora de probabilidade, a última expressão em (3.68) é interpretada de forma que os coeficientes de z^{v_A} indicam a probabilidade da perda ser $X = v_A$. Isto é, existe uma probabilidade de 76,5% da perda da carteira ser zero -

¹ Para outros exemplos com FGP veja Melchiori (2004).

em que nenhum dos devedores entra em *default*. A probabilidade da perda da carteira ser 1 é 13,5%, o que significa que o devedor com exposição igual a 1 ficou inadimplente. Da mesma forma, a probabilidade da perda ser 2 é 8,5%, e acontece quando o devedor com $v_A = 2$ entra em *default*. Por último, existe uma probabilidade de 1,5% que ambos devedores estarão em *default* no fim do período de análise, ocasionando a perda do valor total da carteira.

Outra forma de obter a distribuição de perdas da carteira por meio da FGP é por meio da expressão (2.3) em que

$$P(X = n) = \frac{1}{n!} G_X^{(n)}(0). \quad (3.69)$$

Calculando para $n = 2$ tem-se

$$\begin{aligned} P(X = 2) &= \frac{1}{2!} \frac{d^2}{dz^2} \left(0,765z^0 + 0,135z^1 + 0,085z^2 + 0,015z^3 \right) \Big|_{z=0} \\ &= \frac{1}{2!} (0,17 + 0,09z) \Big|_{z=0} \\ &= 0,085. \end{aligned} \quad (3.70)$$

Recorde que um ajuste nas exposições dos devedores foi feita no início, de forma que a perda real da carteira é dada por múltiplos da unidade de perda U e $\tilde{v}_A = U v_A$. Dessa forma, quando dito que a perda foi de $v_2 = 2$, a exposição original é $\tilde{v}_A = 200 \cdot 2 = 400$.

Como pode ser observado, desde que o número de devedores seja reduzido, este processo pode ser feito rapidamente. Em aplicações reais, o número de devedores em uma carteira de crédito pode compreender até o número total de clientes de uma empresa, e assim, formas mais eficientes de computar a distribuição de perdas foram desenvolvidas.

Como visto na seção 3.3, conclui-se que a aproximação da distribuição de Bernoulli pela Poisson é válida sob algumas condições. Esta aproximação é vantajosa pois a distribuição de Poisson possibilita que a função geradora de probabilidade seja escrita na forma da expressão recursiva, tornando o algoritmo de estimação mais eficiente.

Este exemplo mostrou a teoria inicial por trás do modelo CR+ e como a distribuição das perdas é estimada por funções geradoras de probabilidade. O próximo exemplo avança na teoria apresentada anteriormente, e os conceitos são utilizados em uma carteira para ilustrar o cálculo dos parâmetros necessários para aplicação da expressão recursiva.

3.5.2 Aplicação do Modelo CreditRisk+

Considere uma carteira de crédito hipotética com cinco devedores, sendo que estes estão em um único setor e sem risco idiossincrático. Isto significa que $K = 1$ e $w_{A0} = 0$, $w_{Ak} = 1$ para $A = 1, \dots, 5$. Supondo também que a volatilidade da probabilidade de

TABELA 2 – Exemplo 2 - Dados da carteira

A	\tilde{p}_A	\tilde{v}_A	p_A	v_A
1	0,01	100	0,0100	1
2	0,02	150	0,0150	2
3	0,03	250	0,0250	3
4	0,04	200	0,0400	2
5	0,05	400	0,0500	4

default seja $\sigma_A = p_A/2$ e adotando o número de faixas $F = 4$, a Tabela 2 apresenta os dados de entrada do exemplo.

Observe que \tilde{p}_A e \tilde{v}_A são as probabilidade de *default* e exposição do devedor A . A unidade de perda, ou tamanho da faixa U é determinada por

$$U = \frac{\max [\tilde{v}_A]}{F} = \frac{400}{4} = 100 \quad (3.71)$$

e então, p_A e v_A são as versões ajustadas pelas expressões (3.2) e (3.4) de forma análoga ao calculado para o exemplo da Ilustração 1, apresentado na seção 3.5.1.

Como o número de faixas é menor que a quantidade de devedores na carteira, é necessário agregá-los em suas faixas comuns. De acordo com a Tabela 2, os devedores 2 e 4 ficaram na mesma faixa de exposição, e como visto em (3.5) suas probabilidades de *default* são somadas e reduzidas à uma única faixa. Este processo reduz a quantidade de exposições utilizadas no cálculo, o que pode ser visto na Tabela 3.

TABELA 3 – Exemplo 2 - Dados de entrada do modelo CR+

Faixas	v_A	p_A	σ_A
1	1	0,0100	0,0050
2	2	0,0550	0,0275
3	3	0,0250	0,0125
4	4	0,0500	0,0250

A título de exemplo, nesta seção a estimativa manual será feita apenas para as primeiras iterações do algoritmo recursivo. Os parâmetros necessários para o cálculo da distribuição de perdas serão estimados para a carteira fictícia com o propósito de aplicar o modelo CR+ a seguir.

Como, neste exemplo, os devedores estão todos no mesmo setor, alguns parâmetros precisam ser determinados apenas para $k = 1$. De acordo com (3.20), (4.3), (3.25), (3.28),

e (3.32), tem-se

$$\begin{aligned}
\mu_k &= \sum_A^4 w_{Ak} p_A \Rightarrow \mu_1 = \sum_A^4 w_{A1} p_A = 0,0100 + 0,0550 + 0,0250 + 0,0500 = 0,1400 \\
\sigma_k &= \sum_A^4 w_{Ak} \sigma_A \Rightarrow \sigma_1 = 0,0050 + 0,0275 + 0,0125 + 0,0250 = 0,0700 \\
\alpha_k &= \frac{\mu_k^2}{\sigma_k^2} \Rightarrow \alpha_1 = \frac{(0,14)^2}{(0,07)^2} = 4 \\
\delta_k &= \frac{\mu_k}{\mu_k + \alpha_k} \Rightarrow \delta_1 = \frac{0,1400}{0,1400 + 4} = 0,0338 \\
\eta_k &= \alpha_k \delta_k \Rightarrow \eta_1 = 4 \cdot 0,0338 = 0,1352
\end{aligned}$$

Para a determinação dos coeficientes a_n e b_n ainda é necessário determinar $c_{r,k}$ que, de acordo com (3.45) é dado por

$$c_{r,k} = \sum_{A:v_A=r} w_{Ak} p_A, \quad (3.72)$$

como $w_{Ak} = 1$ para todo A simplifica-se esta expressão por $c_r = \sum_{A:v_A=r} p_A$ e

$$\begin{aligned}
c_0 &= 0 \\
c_1 &= 0,100 \\
c_2 &= 0,055 \\
c_3 &= 0,025 \\
c_4 &= 0,050.
\end{aligned}$$

De acordo com (3.49) Os coeficientes b_n são determinados por

$$\begin{aligned}
b_n &= \sum_{n_1+\dots+n_K=n} \prod_{k=1}^K \left(\delta_{0n_k} - \mathbb{I}(n_k) \frac{\delta_k c_{n_k,k}}{\mu_k} \right) \\
&= \sum_{n_1=n} \left(\delta_{0n_1} - \mathbb{I}(n_1) \frac{\delta_1 c_{n_1}}{\mu_1} \right). \quad (3.73)
\end{aligned}$$

Dessa forma, tem-se

$$\begin{aligned}
b_0 &= \delta_{00} - \mathbb{I}(0) \frac{\delta_1 c_0}{\mu_1} = 1 - 0 = 1 \\
b_1 &= \delta_{01} - \mathbb{I}(1) \frac{\delta_1 c_1}{\mu_1} = 0 - 1 \cdot \frac{0,0338 \cdot 0,01}{0,14} = -0,00242 \\
b_2 &= \delta_{02} - \mathbb{I}(2) \frac{\delta_1 c_2}{\mu_1} = 0 - 1 \cdot \frac{0,0338 \cdot 0,055}{0,14} = -0,01329 \\
b_3 &= \delta_{03} - \mathbb{I}(3) \frac{\delta_1 c_3}{\mu_1} = 0 - 1 \cdot \frac{0,0338 \cdot 0,025}{0,14} = -0,00604 \\
b_4 &= \delta_{04} - \mathbb{I}(4) \frac{\delta_1 c_4}{\mu_1} = 0 - 1 \cdot \frac{0,0338 \cdot 0,005}{0,14} = -0,01208
\end{aligned}$$

De forma análoga, como em (3.50), a expressão para os coeficientes a_n é dada por

$$a_n = \sum_{k=0}^K \sum_{n_0+\dots+n_K=n} \eta_k \frac{c_{n_{k+1},k}}{\mu_k} (n_k + 1) \mathbb{I}(n_k + 1) \prod_{j=0, j \neq k}^K \left(\delta_{0n_j} - \mathbb{I}(n_j) \delta_j \frac{c_{n_j,j}}{\mu_j} \right), \quad (3.74)$$

como, neste exemplo, somente o risco sistemático está presente (e o setor $k = 0$, que representa o risco idiossincrático, não) e considerando apenas um setor, a expressão pode ser escrita como

$$a_n = \sum_{n_1=n} \eta_1 \frac{c_{n_1+1}}{\mu_1} (n_1 + 1) \mathbb{I}(n_1 + 1). \quad (3.75)$$

Logo, os coeficientes do polinômio $A(z)$ são

$$\begin{aligned} a_0 &= \eta_1 \frac{c_1}{\mu_1} (1) \mathbb{I}(1) = 0,1352 \cdot \frac{0,01}{0,14} \cdot 1 \cdot 1 = 0,0096 \\ a_1 &= \eta_1 \frac{c_2}{\mu_1} (2) \mathbb{I}(2) = 0,1352 \cdot \frac{0,055}{0,14} \cdot 2 \cdot 1 = 0,1062 \\ a_2 &= \eta_1 \frac{c_3}{\mu_1} (3) \mathbb{I}(3) = 0,1352 \cdot \frac{0,025}{0,14} \cdot 3 \cdot 1 = 0,0724 \\ a_3 &= \eta_1 \frac{c_4}{\mu_1} (4) \mathbb{I}(4) = 0,1352 \cdot \frac{0,05}{0,14} \cdot 4 \cdot 1 = 0,1932. \end{aligned}$$

Finalmente, a distribuição de perda agregada da carteira é expressa por (3.63)

$$g_{n+1} = \frac{1}{b_0(n+1)} \left(\sum_{j=0}^{\min(d_A, n)} a_j g_{n-j} - \sum_{j=0}^{\min(d_B, n)-1} (n-j) b_{j+1} g_{n-j} \right). \quad (3.76)$$

Os primeiros termos da distribuição são, por (3.55) e (3.56),

$$\begin{aligned} g_0 &= e^{-\mu_0} \prod_{k=1}^K (1 - \delta_k)^{\alpha_k} = (1 - \delta_1)^\alpha = (1 - 0,14)^4 = 0,8714 \\ g_1 &= g_0 \frac{a_0}{b_0} = 0,8714 \cdot \frac{0,0096}{1} = 0,0084. \end{aligned}$$

Para $n = 1$

$$\begin{aligned} g_2 &= \frac{1}{2} \left(\sum_{j=0}^1 a_j g_{1-j} - \sum_{j=0}^0 (1-j) b_{j+1} g_{1-j} \right) \\ &= 0,5 (a_0 g_1 + a_1 g_0 - b_1 g_1) \\ &= 0,5 (0,0096 \cdot 0,0084 + 0,1062 \cdot 0,8714 - (-0,0024 \cdot 0,0084)) \\ &= 0,0464. \end{aligned}$$

Para $n = 2$

$$\begin{aligned} g_3 &= \frac{1}{3} \left(\sum_{j=0}^2 a_j g_{2-j} - \sum_{j=0}^1 (2-j) b_{j+1} g_{2-j} \right) \\ &= 0,33 (a_0 g_2 + a_1 g_1 + a_2 g_0 - (2b_1 g_2 + b_2 g_1)) \\ &= 0,0216. \end{aligned}$$

Para $n = 3$

$$\begin{aligned}
 g_4 &= \frac{1}{4} \left(\sum_{j=0}^3 a_j g_{3-j} - \sum_{j=0}^2 (3-j) b_{j+1} g_{3-j} \right) \\
 &= 0,25 (a_0 g_3 + a_1 g_2 + a_2 g_1 + a_3 g_0 - (3 b_1 g_3 + 2 b_2 g_2 + b_3 g_1)) \\
 &= 0,0439.
 \end{aligned}$$

Este processo é feito até que soma acumulada de g se aproxime de 1, isto é, até $G(j) = g_0 + g_1 + \dots + g_j$ atingir um percentil admissível ou até o número máximo de iterações adotado. Neste exemplo, emprega-se o percentil de 0,9999 e número de iterações $n_{it} = 50$. A Tabela 4 mostra as próximas iterações da estimação da distribuição de perdas para esta ilustração, assim como os coeficientes dos polinômios $A(z)$ e $B(z)$.

TABELA 4 – Exemplo 2 - Iterações do algoritmo

n	a_n	b_n	g_n	$G(n)$
0	0,0096	1,0000	0,8714	0,8714
1	0,1062	-0,0024	0,0084	0,8799
2	0,0724	-0,0132	0,0464	0,9262
3	0,1932	-0,0060	0,0216	0,9478
4		-0,0121	0,0439	0,9917
5			0,0019	0,9937
6			0,0032	0,9968
7			0,0014	0,9982
8			0,0014	0,9996
9			0,0001	0,9998
10			0,0001	0,9999
11			0,0001	0,9999

Observe que $\sum_A v_A = 1 + 2 + 3 + 4 = 10$ é a perda máxima da carteira, e acontece quando todos os devedores entram em estado de *default*. É intuitivo então que a distribuição de perdas seja calculada até $n = 10$, mas por conta das aproximações realizadas no modelo, na prática calcula-se até $n = 11$.

Para o cálculo do VaR, segundo a expressão (2.16), pode-se fazer a seguinte interpolação linear entre os dois pontos q_1 e q_2

$$VaR_\epsilon = U \left[(v_{A_{q_2}} - 1) + \left(\frac{\epsilon - q_1}{q_2 - q_1} \right) \right]$$

em que $\epsilon \in [q_1, q_2]$ é o nível de confiança adotado para estimação do VaR, e $v_{A_{q_2}}$ representa o valor de perda discretizada associada à probabilidade acumulada da distribuição de perda agregada até q_2 .

Considere $\epsilon = 95\%$ para esta ilustração, conseqüentemente tem-se $q_1 = 0,9478$, $q_2 = 0,9917$ e $v_{A_{q_2}} = 4$, pela Tabela 4,. O tamanho da faixa é $U = 100$ conforme (3.71), e

a estimação do VaR é feita por

$$VaR_{0,95} = 100 \left[(4 - 1) + \left(\frac{0,95 - 0,9478}{0,9917 - 0,9478} \right) \right] = 304,94. \quad (3.77)$$

A interpretação de $VaR = 304,94$ pode ser entendida como a existência de uma chance de apenas 5% de que a perda da carteira será superior a 304,94 unidades monetárias, ou de outra forma, tendo uma probabilidade de 95% de que a perda não ultrapassará 304,94 unidades monetárias.

4 DADOS E MÉTODOS

Considerando os conceitos apresentados até o momento, este capítulo se dedica a introduzir a parte de aplicação no presente trabalho. Assim, serão apresentados os dados utilizados para estimação do risco por meio do modelo CreditRisk+, a forma de obtenção dos dados de entrada do modelo, assim como as técnicas empregadas para tanto.

Relembrando, do capítulo anterior, os dados de entrada do modelo são: a exposição de cada devedor, a probabilidade de *default*, a volatilidade da probabilidade de *default*, além da matriz de pesos w_{Ak} dos diferentes setores considerados. A exposição dos devedores e sua participação em cada setor normalmente são dados já disponíveis para cada análise de risco, sendo fornecidas pela instituição interessada, restando assim a estimação da probabilidade de *default* e sua volatilidade.

Dessa forma, a primeira parte deste capítulo apresenta os dados utilizados a fim de esclarecer a natureza das carteiras analisadas. A estimação do risco de crédito ocorrerá para duas carteiras distintas. A primeira diz respeito à uma carteira de crédito real de uma empresa securitizadora. A segunda está relacionada aos dados de crédito de uma instituição financeira alemã. O restante do capítulo mostrará como os dados de entrada do modelo CreditRisk+ foram obtidos e preparados.

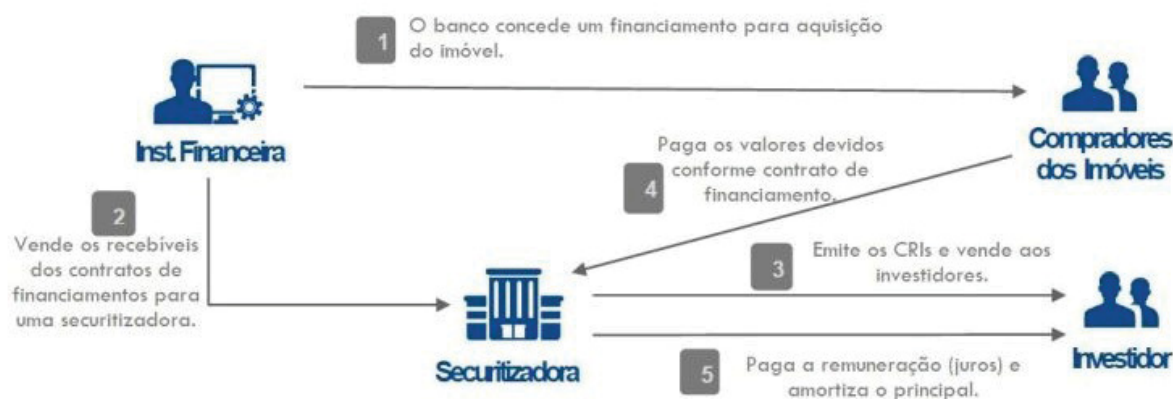
4.1 CARTEIRA DE CRÉDITO DE SECURITIZAÇÃO

No presente trabalho, a primeira carteira de crédito avaliada está relacionada a uma empresa securitizadora. A securitização é uma operação financeira de captação de recursos que transforma uma dívida qualquer com determinado devedor em títulos negociáveis no mercado de capitais. Neste caso, a empresa securitizadora adquire uma dívida imobiliária, e emite títulos de crédito conhecidos como Certificado de Recebíveis Imobiliários (CRI).

A Figura 2 mostra o processo de securitização para a emissão do CRI. As partes envolvidas nas etapas da securitização são o originador, a empresa securitizadora e os investidores. O originador pode ser uma instituição financeira que, em um primeiro momento, concede financiamentos ou empréstimos para os compradores dos imóveis. A instituição financeira pode então vender tal dívida imobiliária, na forma dos recebíveis dos contratos de financiamentos para a empresa securitizadora. Dessa forma, os compradores dos imóveis passam a estar em dívida com a securitizadora em vez do originador. Por fim, a empresa de securitização emite os CRIs, vendendo para investidores interessados. Dessa forma, o processo de securitização é uma “transformação” de títulos e contratos representativos de crédito, com difícil circulação, tornando-os em um instrumento mais

facilmente negociável no mercado de capitais (AVELINO, 2014).

FIGURA 2 – Processo de Securitização



Fonte: Peixoto (2016)

A securitização se refere principalmente à monetização de dívidas, desse modo, os riscos envolvidos no processo estão relacionados principalmente ao pagamento da obrigação e não ao desempenho de uma entidade externa em particular (FABOZZI, 2008). Para a carteira de crédito em questão, os compradores dos imóveis são os devedores, e o risco de crédito surge no momento em que a empresa adquire uma carteira de financiamentos imobiliários e é então responsável por sua gestão.

Esta carteira de crédito também foi utilizada em Oliveira (2018) para a avaliação do risco de crédito. Neste trabalho, métodos de *Credit Scoring* foram empregados na classificação dos devedores entre bons e maus pagadores, a fim de oferecer à empresa securitizadora um indicador de quais clientes poderiam entrar em *default*. As técnicas empregadas para a classificação foram a regressão logística e *support vector machine* (máquina de vetores suporte), em que a comparação mostrou maior taxa de discriminação pelo modelo *logit*.

Os dados cedidos são referentes à uma carteira de crédito de financiamentos imobiliários com 515 devedores, no período de junho de 2013 a abril de 2018. Dos 515 devedores presentes na carteira, 165 são considerados inadimplentes e 350 são adimplentes, em uma proporção de 32,02% do estado de inadimplência em relação ao número total de devedores. A base de dados compreende informações cadastrais de cada devedor, do imóvel adquirido pelo mesmo e situação da dívida - se existem parcelas em atraso ou não.

Dada a natureza do processo de securitização, a empresa em questão adquire a carteira com os devedores, não sendo responsável pelo seu cadastro inicial. Dessa forma, os dados disponíveis dos devedores dependem da instituição financeira que vendeu os contratos de financiamentos para a securitizadora, limitando o controle sob as variáveis.

As 23 variáveis explicativas disponíveis eram em sua maioria contratuais ou construídas com base em informações de pagamentos. Dentre outras que foram desconsideradas por serem variáveis resposta, duplicadas ou altamente correlacionadas, os atributos considerados para análise final foram resumidos em 12 variáveis. Dentre estas características, 9 são variáveis quantitativas e 3 são qualitativas. A tabela 5 apresenta um resumo dos atributos presentes na base de dados.

TABELA 5 – Carteira 1 - Variáveis explicativas da carteira

Variável	Descrição	Tipo
Area_Priv	Área do lote	Numérica
Idade	Idade do solicitante	Numérica
Casado	Estado civil do solicitante	Categórica
N_Meses_Assinatura	Número de meses entre a data atual e a assinatura do contrato.	Numérica
N_Meses_Adimp	Número de meses em que o solicitante pagou sem atrasos.	Numérica
Valor_Rec_Atualiz	Valor recebido atualizado.	Numérica
LTV	Saldo para quitação dividido pelo valor de compra atualizado.	Numérica
Preco_Tabela	Preço de tabela dos lotes do condomínio.	Numérica
p_Garantia	Percentual dado em garantia.	Numérica
Cessao	Se houve cessão ou não	Categórica
Renegociacao	Se houve renegociação	Categórica
Meses_no_contrato	Número de meses no contrato	Numérica

As características pessoais dos devedores são representadas pelas variáveis Idade e Casado, enquanto o imóvel adquirido é descrito pelas variáveis Area_Priv e Preco_Tabela. O restante dos atributos se referem à situação da dívida dos devedores, como por exemplo a variável N_Meses_Adimp que indica a quantidade de vezes que o devedor pagou uma parcela do financiamento com nenhum dia em atraso. Já a variável LTV indica o *Loan to Value*, uma proporção entre o valor devido em relação ao total da dívida. A variável Cessao é categórica e indica se o financiamento já passou por um processo de cessão, ou seja, se um primeiro devedor transferiu sua posição na relação obrigacional para um terceiro, que assume o pagamento e posse do imóvel. A variável Renegociacao também é binária, indicando se já houve em algum momento a renegociação de pagamentos atrasados para o devedor atual. Essas 12 variáveis são atribuídas para cada um dos 515 devedores presentes na carteira de crédito.

A variável resposta é binária e indica se cada devedor é inadimplente ou não. Para

a classificação, foi considerado como inadimplente qualquer devedor com uma ou mais parcelas em aberto que tenham mais de 90 dias de atraso.

Para a estimação da perda pelo modelo CreditRisk+, o primeiro dado de entrada é a exposição individual de cada devedor. Para esta carteira, a exposição é estimada com base no valor devido à empresa securitizadora no período de um ano. Ou seja, a exposição é a soma das próximas 12 parcelas do financiamento. Dessa forma, supõe-se que dada a inadimplência, a empresa perderá o valor restante devido em um ano.

Segundo Balzarotti, Falkenheim e Powell (2002), a maioria dos analistas considera este horizonte de tempo como apropriado para medir a perda de capital, visto que um ano é suficiente para as instituições financeiras tomarem medidas preventivas de risco ou recolher um nível de capital que compensaria tal perda caso viesse a ocorrer.

4.2 CARTEIRA DE CRÉDITO ALEMÃ

A segunda base de dados utilizada representa um conjunto de dados de crédito para uma instituição financeira alemã, disponível em Dheeru e Taniskidou (2018). A carteira de crédito compreende 1000 devedores, dos quais 300 são considerados em estado de inadimplência e 700 adimplentes. A base de dados inclui informações diversas sobre cada devedor, desde sua idade, valor do empréstimo e se possui outros investimentos.

No total, 20 atributos estão disponíveis para avaliação do risco, entre estas características, 7 são variáveis quantitativas e 13 são qualitativas. Na Tabela 6, um resumo dos atributos presentes na base de dados é apresentado, juntamente com uma breve descrição das variáveis.

Enquanto a carteira apresentada na seção anterior tinha suas variáveis categóricas binárias, possuindo apenas dois níveis (sim e não), nesta carteira alguns atributos possuem diversas categorias. Neste caso, é necessário codificar as variáveis categóricas, o que pode ser feito por meio de variáveis *dummies*¹. Tomando o atributo *Personal* como exemplo, o sexo e o estado civil do solicitante foram organizadas em cinco categorias:

- a) P1: masculino e divorciado;
- b) P2: feminino e divorciada/casada;
- c) P3: masculino e solteiro;
- d) P4: masculino e casado/viúvo;
- e) P4: feminino e solteira.

Se cada categoria fosse codificada com números de 1 a 5, é possível perceber que não existe uma relação de aumento entre esses grupos. De fato, ser uma mulher

¹ Uma variável *dummy* ou variável indicadora é uma variável artificial binária criada para representar um atributo com dois ou mais categorias/níveis distintos.

TABELA 6 – Carteira 2 - Variáveis explicativas da carteira

Variável	Descrição	Tipo
<i>CheckingAccountStatus</i>	Status da conta do solicitante (salário declarado)	Categórica
<i>Duration</i>	Duração do empréstimo solicitado (mês)	Numérica
<i>CreditHistory</i>	Histórico de crédito	Categórica
<i>Purpose</i>	Finalidade do crédito solicitado	Numérica
<i>Amount</i>	Valor do crédito solicitado	Numérica
<i>SavingsAccountBonds</i>	Economias do solicitante	Categórica
<i>EmploymentDuration</i>	Tempo no atual emprego	Categórica
<i>InstallmentRatePercentage</i>	Proporção da parcela do empréstimo pela renda do solicitante	Numérica
<i>Personal</i>	Sexo e estado civil	Categórica
<i>OtherDebtorsGuarantors</i>	Outras dívidas/ fiador	Categórica
<i>ResidenceDuration</i>	Tempo na residência atual	Numérica
<i>Property</i>	Outras propriedades	Categórica
<i>Age</i>	Idade	Numérica
<i>OtherInstallmentPlans</i>	Outros parcelamentos	Categórica
<i>Housing</i>	Tipo de moradia	Categórica
<i>NumberExistingCredits</i>	Número de créditos existentes neste banco	Numérica
<i>Job</i>	Profissão	Categórica
<i>NumberPeopleMaintenance</i>	Número de dependentes	Numérica
<i>Telephone</i>	Telefone próprio	Categórica
<i>ForeignWorker</i>	Solicitante é estrangeiro	Categórica

divorciada não é maior do que um homem divorciado, da mesma forma que um homem casado não representa o dobro de uma mulher divorciada. Neste caso, utilizar variáveis *dummy* apresenta uma codificação mais apropriada. Para tanto, as diferentes categorias do atributo são separadas em diferentes variáveis binárias, com valor um (1) indicando a presença da característica e zero representando a ausência.

A Tabela 7 mostra uma ilustração da codificação do atributo *Personal*. Para os cinco níveis de características desta variável, são criadas quatro variáveis *dummy*. A última categoria (feminino e solteira) não é representada por uma variável binária própria, mas pela sua ausência nas outras quatro variáveis, o que pode ser observado para o Devedor 5 na tabela.

TABELA 7 – Carteira 2 - Codificação *dummy* para o atributo *Personal*

Devedor	<i>Personal</i>	Variáveis <i>dummy</i>			
		<i>male.divorced</i>	<i>female.notsingle</i>	<i>male.single</i>	<i>male.married</i>
1	P2	0	1	0	0
2	P1	1	0	0	0
3	P3	0	0	1	0
4	P4	0	0	0	1
5	P5	0	0	0	0

Assim, todas as variáveis categóricas com mais de dois níveis podem ser codificadas dessa forma. Dado que a maior parte das variáveis da base de dados é composta de atributos categóricos (13 das 20), ao fim do processo de codificação *dummy*, o número de variáveis do problema passa de 20 características para 52.

A variável resposta é binária e indica se cada cliente é inadimplente ou não. Para esta carteira, tal atributo é descrito pela variável *Class* e possui valor um para devedores inadimplentes e zero para devedores adimplentes.

Para estes dados o valor em exposição será o valor de crédito concedido a cada devedor, ou seja, a variável *Amount* na Tabela 6.

Cabe lembrar que para a implementação do modelo CR+, os dados de entrada exigidos compreendem ainda a probabilidade de *default* individual e a volatilidade da probabilidade de *default* de cada devedor. Estes dois últimos dados são o assunto de discussão das próximas seções.

4.3 ESTIMAÇÃO DA PROBABILIDADE DE *DEFAULT*

Dada a natureza do modelo CreditRisk+, pode-se dizer que seu dado de entrada principal é a probabilidade de *default*. Conforme visto no capítulo 2, os modelos de *Credit Scoring* são utilizados geralmente para classificar e discriminar entre clientes “bons” e “maus”, e nesta seção utiliza-se tais técnicas para estimar a probabilidade de serem “maus”, em vez de apenas classificá-los dessa forma.

De acordo com as características apresentadas na seção 2.4, o método adotado para a estimação da probabilidade de *default* neste trabalho é a regressão logística, sendo indicada para casos em que a variável resposta do modelo é dicotômica, assumindo apenas dois valores possíveis. Para a variável dependente, considera-se que o evento de interesse é o *default*, representado pelo valor um, enquanto a adimplência é descrita pelo valor zero.

De forma a avaliar o modelo e sua capacidade de classificação, os dados foram separados em conjunto de treinamento e teste. Dessa forma, 80% dos dados foram utilizados para a construção do modelo, enquanto os 20% restantes foram utilizados para avaliação do modelo encontrado.

Para a seleção das variáveis foi utilizado o método *stepwise*. Esta técnica baseia-se em um algoritmo estatístico que verifica a “importância” das variáveis independentes, incluindo ou excluindo-as do modelo com base em um critério de decisão (HOSMER; LEMESHOW, 2000). Nesta análise, a medida utilizada para a escolha das variáveis é o critério de informação de Akaike (AIC), descrita na equação (4.1) como

$$AIC = 2k - 2\ln(\hat{L}) \quad (4.1)$$

em que \hat{L} é o valor máximo da função de verossimilhança para o modelo e k é o número de parâmetros estimados.

O AIC é uma medida que recompensa a qualidade do ajuste - avaliado pela função de verossimilhança - e inclui uma penalidade para o número de parâmetros estimados k . Dessa forma, o modelo com o menor valor de AIC é favorecido (BURNHAM; ANDERSON, 2002). Este critério pode ser utilizado na seleção de variáveis, pois dado um conjunto de modelos, o AIC estima a performance de cada modelo em relação aos outros. No método *stepwise*, diferentes modelos são estimados ao incluir ou excluir variáveis, sendo assim, adota-se o AIC.

4.3.1 Carteira de Securitização

Utilizando a função *glm* do software de computação estatística R (R CORE TEAM, 2018), estima-se um modelo *logit* com todas as 12 variáveis independentes disponíveis na análise. A partir deste modelo completo, a função *stepAIC* do pacote MASS (VENABLES; RIPLEY, 2002) é utilizada para a escolha das variáveis mais importantes conforme o AIC.

O modelo final foi composto pelo intercepto e cinco variáveis explicativas. As variáveis incluídas no modelo foram Idade (x_2), N_Meses_Assinatura (x_4), N_Meses_Adimp (x_5), Valor_Rec_Atualiz (x_6) e Cessao (x_{10}). A expressão matemática do modelo é escrita como (4.2) :

$$\ln\left(\frac{p}{1-p}\right) = -0,544 - 0,041 x_2 + 0,085 x_4 - 2,748 x_5 - 1,99 \cdot 10^{-5} x_6 - 1,772 x_{10} . \quad (4.2)$$

As variáveis x_2 , x_5 , x_6 e x_{10} possuem coeficientes negativos, o que significa que, quanto maiores forem os valores assumidos por essas variáveis, menor a probabilidade de *default*. Esta interpretação é bem intuitiva e esperada para as variáveis x_2 , x_5 e x_6 , em que quanto maior a idade, o número de meses pagos sem atrasos e o valor recebido do financiamento, indicam um devedor com menor probabilidade de *default*. Por meio do desenvolvimento deste modelo, é possível concluir também que os devedores que adquiriram o financiamento por meio de um processo de cessão (x_{10}) possuem menor probabilidade de entrar em inadimplência, como indica o coeficiente negativo.

A variável x_4 apresentou coeficiente positivo, mostrando que quanto maior o número de meses decorrido entre a assinatura do contrato e a data atual, maior a probabilidade de um devedor ser inadimplente.

Conforme Vojtek e Kočenda (2006) uma das fragilidades da regressão logística é o fato da técnica ser sensível a alta correlação entre variáveis, neste caso, deve-se garantir que não existam tais relações, ou aplicar outros métodos de regularização. Segundo Brito e Assaf Neto (2008), a inclusão de variáveis altamente correlacionadas pode prejudicar a capacidade preditiva do modelo, no sentido de que tais variáveis fornecem informações semelhantes na explicação da variável resposta. Dessa forma, na análise exploratória dos dados, as 23 variáveis iniciais disponíveis foram reduzidas a 12 de acordo com coeficientes de correlação e também por meio da interpretação de tais variáveis, de forma que se algumas delas possuísem significado semelhante a outras, seriam assim desconsideradas. A Tabela 8 apresenta também os coeficientes de correlação de Pearson das variáveis do modelo final, demonstrando que não são altamente correlacionadas.

TABELA 8 – Carteira 1 - Matriz de correlação das variáveis explicativas do modelo

	x_2	x_4	x_5	x_6	x_{10}
x_2	1				
x_4	0,2781	1			
x_5	0,1073	0,0342	1		
x_6	0,2000	0,4421	0,1884	1	
x_{10}	-0,0133	0,1757	-0,0204	0,1273	1

Uma forma de validação do modelo é feita pela análise da significância estatística de cada variável. Ou seja, deseja-se avaliar se uma variável é importante para o modelo, ou se possui um relacionamento significativo com a variável resposta. O teste Wald avalia a hipótese nula de que o parâmetro β estimado é igual a zero. Em um teste de hipótese, assume-se que a hipótese nula é verdadeira, buscando evidências nos dados para rejeitá-la. Se o teste falhar em rejeitar a hipótese nula, significa que a variável não é significativa, e retirá-la do modelo não prejudicará a qualidade de ajuste.

A Tabela 9 mostra os coeficientes do modelo de regressão para cada variável, juntamente com as informações de ajuste como o erro padrão, valor z e valor p.

Neste contexto, o valor z é calculado como resultado da padronização dos coeficientes de regressão logística. Especificamente, o valor z é estimado como o quociente do coeficiente de regressão β pelo seu erro padrão. Este é então utilizado como a estatística de teste para os testes de hipótese de que os coeficientes β encontrados na regressão assumem valor zero.

O valor p é definido como a probabilidade de obter um resultado igual ou mais extremo na direção da hipótese alternativa do que o observado, considerando que a hipótese

nula é verdadeira. Dessa forma, a hipótese nula será rejeitada se o valor p for menor ou igual a um certo valor definido como o nível de significância α .

Adotando o nível de significância de 5%, tem-se que o valor p de todas variáveis é menor que 0,05 e, portanto, rejeita-se a hipótese nula. Concluindo assim que as variáveis independentes selecionadas são estatisticamente diferentes de zero.

TABELA 9 – Carteira 1 - Coeficientes do modelo de regressão

Variável	Coeficiente	Erro Pad.	Valor z	Pr(> z)
Intercepto	-0,544	0,673	-0,808	0,419
x_2	-0,041	0,015	-2,648	0,008
x_4	0,0854	0,012	7,059	$1,68 \times 10^{-12}$
x_5	-2,748	0,455	-6,035	$1,59 \times 10^{-9}$
x_6	$-1,99 \times 10^{-5}$	$7,76 \times 10^{-6}$	-2,569	0,010
x_{10}	-1,772	0,666	-2,659	0,007

Com o modelo definido, os dados do conjunto de teste são utilizados para avaliar a capacidade de estimação do modelo *logit*. Como estas observações não foram usadas na construção do modelo, o percentual de acerto no conjunto de testes é utilizada como forma de avaliar como o modelo está classificando os devedores. A probabilidade de *default* para os dados de teste são calculados por meio da expressão (2.12). O ponto de corte do modelo é 0,5, portanto, os clientes com resultado superior serão classificados como inadimplentes e os com percentual inferior são adimplentes. A Tabela 10 mostra a taxa de acerto do modelo na classificação.

TABELA 10 – Carteira 1 - Matriz de classificação do modelo de regressão

Estimado	Observado		Total	Classificações Corretas
	<i>N-Default</i>	<i>Default</i>		
<i>N-Default</i>	67	6	73	91,78%
<i>Default</i>	5	25	30	83,33%
Total	72	31	103	89,32%

Conforme mostra a Tabela 10, o nível geral de acerto do modelo *logit* é 89,32%, com 92 devedores classificados corretamente dentre os 103 do conjunto de testes. Entre os devedores considerados adimplentes, o modelo classificou corretamente 67 dos 73, com uma taxa de acerto de 91,78%. Para os clientes em *default*, o percentual de acerto foi de 83,33% sendo que o modelo discriminou 25 devedores corretamente e 5 erroneamente.

Outra forma de avaliar a performance do modelo é dada pela área sob a curva ROC (*Receiver Operating Characteristic*). Geometricamente, a curva ROC é um gráfico de eixos (1-especificidade) e sensibilidade, em que tais medidas são calculadas para todas as observações da amostra, considerando diferentes pontos de corte do modelo. A sensibilidade, conhecida também como taxa de verdadeiros positivos, é o percentual de acerto na previsão do evento de interesse quando ele de fato ocorreu. A especificidade, ou taxa de verdadeiros

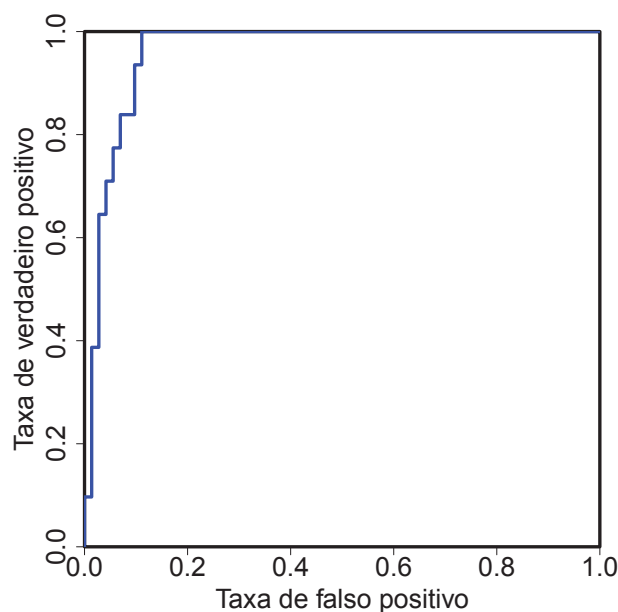
negativos, é a proporção de acerto na previsão da não ocorrência do evento de interesse quando ele de fato não ocorreu.

A área abaixo da curva ROC varia entre zero e um, e mede a capacidade do modelo de discriminar entre as estimações que tiveram a classificação desejada das que não. Hosmer e Lemeshow (2000) apresentam uma regra geral para a interpretação deste valor:

- a) Se $ROC = 0,5$, não há discriminação;
- b) Se $0,7 < ROC < 0,8$ tem-se uma discriminação aceitável;
- c) Se $0,8 < ROC < 0,9$ tem-se uma discriminação excelente;
- d) Se $ROC > 0,9$ tem-se uma discriminação excepcional.

Para a construção da curva ROC e estimação da área sob a curva, as funções presentes no pacote ROCR (SING et al., 2005) foram utilizadas. A Figura 3 mostra a curva ROC do modelo *logit* desenvolvido para a carteira de crédito, revelando que a área sob a curva ROC é de 0,961. Segundo a regra geral apresentada anteriormente, este valor indica uma capacidade excepcional de discriminação do modelo.

FIGURA 3 – Carteira 1 - Curva ROC



Fonte: A autora (2018)

Portanto, o modelo desenvolvido é válido para a estimação das probabilidades de *default* e calcula-se a PD para os 515 devedores da carteira de crédito por meio do modelo *logit* encontrado. A Tabela 11 mostra como a PD está distribuída entre os 515 devedores da carteira.

TABELA 11 – Carteira 1 - Distribuição da Probabilidade de *default*

PD	Devedores	Proporção
$p_A \leq 0,1$	226	43,88%
$0,1 < p_A \leq 0,2$	51	9,90%
$0,2 < p_A \leq 0,3$	30	5,83%
$0,3 < p_A \leq 0,4$	18	3,50%
$0,4 < p_A \leq 0,5$	13	2,52%
$0,5 < p_A \leq 0,6$	20	3,88%
$0,6 < p_A \leq 0,7$	34	6,60%
$0,7 < p_A \leq 0,8$	60	11,65%
$0,8 < p_A \leq 0,9$	62	12,04%
$p_A > 0,9$	1	0,19%

4.3.2 Carteira Alemã

Esta seção apresenta o desenvolvimento do modelo *logit* para a segunda carteira de empréstimos. Todo o processo elaborado para a carteira de crédito de securitização foi reproduzido da forma análoga nesta seção, assim, é possível sintetizar o desenvolvimento e apresentar apenas os resultados.

Do mesmo modo ao realizado para a carteira de crédito anterior, definem-se as variáveis do modelo de regressão logística por meio do método *stepwise*, utilizando como critério a medida de AIC.

No primeiro modelo encontrado pela regressão *stepwise*, algumas variáveis apresentaram valor p maior que o nível de significância de 5%. Pelo teste Wald, isto significa que os coeficientes dessas variáveis são estatisticamente iguais a zero, e sua remoção não deve afetar a performance do modelo. Após desconsiderar algumas variáveis e realizar a regressão logística novamente, o modelo final foi constituído pelo intercepto e 21 variáveis explicativas. A Tabela 12 mostra os coeficientes do modelo de regressão para cada variável, assim como as informações de ajuste como o erro padrão, valor z e valor p.

Observando a última coluna da Tabela 12, com nível de significância de 5%, rejeita-se a hipótese nula e portanto, todas as variáveis independentes do modelo são estatisticamente diferentes de zero.

Para a avaliação da capacidade de previsão do modelo, os dados do conjunto de teste são utilizados para a classificação do modelo. A Tabela 13 mostra a taxa de acerto do modelo na classificação.

Conforme mostra a Tabela 13, o nível geral de acerto do modelo *logit* é 72,50%, com 145 devedores classificados corretamente dentre os 200 do conjunto de testes. Entre os devedores considerados adimplentes, o modelo classificou corretamente 125 dos 156, com uma taxa de acerto de 80,13%. Para os clientes em *default*, o modelo discriminou 20 devedores corretamente e 31 erroneamente, com um percentual de acerto de 45,45%.

TABELA 12 – Carteira 2 - Matriz de Coeficientes do modelo de regressão

Variável	Coeficiente	Erro Pad.	Valor z	Pr(> z)
Intercepto	-4,563	0,672	-6,794	$1,08 \times 10^{-11}$
<i>Duration</i>	0,028	0,010	2,901	0,004
<i>Amount</i>	0,000	0,000	2,809	0,005
<i>InstallmentRatePercentage</i>	0,276	0,093	2,974	0,003
<i>Age</i>	-0,019	0,009	-2,169	0,030
<i>CheckingAccountStatus.lt.0</i>	1,376	0,234	5,888	$3,09 \times 10^{-9}$
<i>CheckingAccountStatus.0.to.200</i>	1,126	0,229	4,922	$8,55 \times 10^{-7}$
<i>CreditHistory.NoCredit.AllPaid</i>	1,190	0,474	2,513	0,012
<i>CreditHistory.ThisBank.AllPaid</i>	1,930	0,431	4,483	$7,35 \times 10^{-6}$
<i>CreditHistory.PaidDuly</i>	0,421	0,205	2,054	0,040
<i>Purpose.NewCar</i>	1,797	0,357	5,035	$4,78 \times 10^{-7}$
<i>Purpose.Furniture.Equipment</i>	1,053	0,370	2,842	0,004
<i>Purpose.Radio.Television</i>	0,760	0,371	2,050	0,040
<i>Purpose.DomesticAppliance</i>	2,296	0,897	2,561	0,010
<i>Purpose.Repairs</i>	1,789	0,632	2,831	0,005
<i>Purpose.Education</i>	1,784	0,471	3,788	$1,51 \times 10^{-4}$
<i>Purpose.Business</i>	1,344	0,436	3,084	0,002
<i>SavingsAccountBonds.lt.100</i>	0,992	0,208	4,767	$1,86 \times 10^{-6}$
<i>EmploymentDuration.4.to.7</i>	-0,749	0,270	-2,774	0,006
<i>Personal.Male.Single</i>	-0,443	0,196	-2,265	0,023
<i>Property.RealEstate</i>	-0,565	0,224	-2,519	0,012
<i>Housing.Rent</i>	0,531	0,242	2,197	0,028

TABELA 13 – Carteira 2 - Matriz de classificação do modelo de regressão

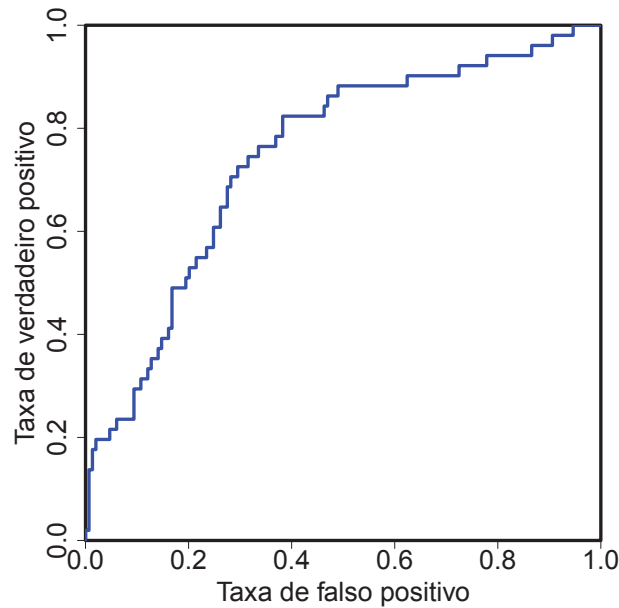
Estimado	Observado		Total	Classificações Corretas
	<i>N-Default</i>	<i>Default</i>		
<i>N-Default</i>	125	31	156	80,13%
<i>Default</i>	24	20	44	45,45%
Total	149	51	200	72,50%

Considerando o nível de acerto de 72,50%, o atual modelo *logit* parece ser aceitável. Contudo, observando que a taxa de acerto foi de 45,45% para os devedores em *default*, o modelo possui melhor capacidade de discriminação para clientes adimplentes.

A Figura 4 mostra a curva ROC referente às previsões deste modelo. Neste caso, a área sob a curva ROC encontrada é 0,745, o que indica uma discriminação aceitável segundo o critério de Hosmer e Lemeshow (2000).

Portanto, o modelo desenvolvido é válido para a estimação das probabilidades de *default*, mas possui fraca capacidade de classificar os devedores considerados inadimplentes. Dada a natureza da base dados, em que o número de atributos é superior à carteira de securitização, esta performance seria um indicativo de que outros métodos podem ser mais apropriados para esta base de dados em vez da regressão logística. Uma vez que o objetivo

FIGURA 4 – Carteira 2 - Curva ROC



Fonte: A autora (2018)

deste trabalho não é comparar modelos de *Credit Scoring*, mas estimar a distribuição de perdas da carteira em questão, a estimação da probabilidade de *default* dos devedores é considerada uma das etapas de tal processo. Assim, considera-se os resultados obtidos por este modelo *logit*, com a ressalva de que não houve comparação com outras técnicas de estimação.

A Tabela 14 mostra como a probabilidade de *default* está distribuída entre os 1000 devedores da carteira.

TABELA 14 – Carteira 2 - Distribuição da Probabilidade de *default*

PD	Devedores	Proporção
$p_A \leq 0,1$	257	25,70%
$0,1 < p_A \leq 0,2$	192	19,20%
$0,2 < p_A \leq 0,3$	136	13,60%
$0,3 < p_A \leq 0,4$	90	9,00%
$0,4 < p_A \leq 0,5$	88	8,80%
$0,5 < p_A \leq 0,6$	77	7,70%
$0,6 < p_A \leq 0,7$	63	6,30%
$0,7 < p_A \leq 0,8$	43	4,30%
$0,8 < p_A \leq 0,9$	40	4,00%
$p_A > 0,9$	14	1,40%

4.4 ESTIMAÇÃO DA VOLATILIDADE DA TAXA DE *DEFAULT*

Um parâmetro essencial na estimação de perdas pelo modelo CR+ é σ_k^2 em (3.24). Tal parâmetro é conhecido como a volatilidade das taxas de *default* de um setor k e mede

o grau a qual a probabilidade de *default* na carteira varia em relação à média μ_k . Segundo Balzarotti, Falkenheim e Powell (2002), sua interpretação no modelo é entendida como a sensibilidade às mudanças na economia ou em outros fatores de risco que influenciam a carteira de crédito. Ou seja, um valor de variância σ_k^2 alta significa que a probabilidade de *default* será alta em momentos desfavoráveis, e uma σ_k^2 baixa representa que a capacidade de pagamento dos devedores não é muito afetada por tais mudanças.

Para estimar a volatilidade de cada setor, CSFP (1997) associa à probabilidade de *default* p_A um desvio padrão σ_A para cada devedor dentro do setor. Dessa forma, a volatilidade σ_k para S_k , $k = 1, \dots, K$ pode ser estimada como

$$\sigma_k = \sum_A w_{Ak} \sigma_A. \quad (4.3)$$

Kluge e Lehrbass (2004) também apresentam uma forma alternativa para a estimação de σ_k na forma

$$\sigma_k = \sum_A \sqrt{w_{Ak}} \sigma_A. \quad (4.4)$$

Em situações ideais é possível estimar a volatilidade da probabilidade de *default* dos devedores para aplicação do modelo, mas em casos onde geralmente tal tarefa é inviável ou de complicada execução, sugestões para a estimação deste parâmetro foram feitas.

De acordo com CSFP (1997), a relação σ_A/p_A é geralmente de ordem 1, o que significa que $\sigma_A = p_A$. Contudo, para carteiras específicas em que a ausência de dados não permite a estimação do parâmetro real, uma proporção única e comum a todos os devedores pode ser adotada. Dessa forma, define-se a relação $\sigma_A = \omega p_A$, onde ω representa a proporção entre a média e o desvio padrão da probabilidade de *default* do devedor A . A volatilidade dos setores pode ser então estimada como

$$\sigma_k = \sum_A w_{Ak} \sigma_A = \sum_A w_{Ak} (\omega p_A). \quad (4.5)$$

Nos estudos de Balzarotti, Falkenheim e Powell (2002), o modelo CR+ foi utilizado para a estimação do capital regulatório do Banco Central da Argentina. Os dados históricos de poucos anos não permitiram a estimação da proporção, em que então, adotou-se $\omega = 0.3$ e $\omega = 0.5$ de acordo com dados externos. Em Schechtman et al. (2004), empregou-se o CR+ para estimação do capital regulatório de bancos brasileiros, onde diversos valores, entre 20% e 130% são testados para a taxa de volatilidade ω . Neste caso, concluiu-se que $\omega = 110\%$ produziu os melhores resultados quando compara-se o modelo CR+ e o cálculo do capital regulatório dos bancos. Já CSFP (1997), sugeriu utilizar a volatilidade $\omega = 50\%$, demonstrando em sua ilustração da aplicação do modelo.

Dessa forma, este trabalho tem o propósito de analisar o impacto do parâmetro ω e consequentemente da volatilidade da taxa de *default* dos devedores. Diferentes valores para o parâmetro serão utilizados para a estimação do risco da carteira em questão.

5 RESULTADOS NUMÉRICOS

O presente capítulo apresenta os resultados obtidos na implementação do modelo CreditRisk+. Conforme visto no capítulo 3, o resultado do algoritmo recursivo do modelo CR+ é a distribuição de probabilidade de perda da carteira de crédito. De posse da distribuição de perdas, a estimação do risco é feita por meio da métrica Valor em Risco e também pelo Valor em Risco Condicional. Neste capítulo, os resultados da estimação do risco são apresentados para as duas carteiras analisadas neste trabalho.

A primeira carteira compreende a base de dados proveniente de uma empresa securitizadora, enquanto a segunda carteira se refere a uma base de dados de uma instituição financeira alemã. Estas serão denominadas de carteiras 1 e 2, respectivamente, no decorrer do capítulo.

Caso não seja mencionado, assume-se que o modelo padrão aplicado terá o número de faixas $F = 100$ e a volatilidade da taxa de *default* é dada por $\omega = 0.5$, conforme sugestões de CSFP (1997) e Balzarotti, Falkenheim e Powell (2002). Além disso, adota-se um único setor sem risco idiossincrático. Admitir apenas um setor significa que as probabilidades de *default* de todos os devedores são influenciados por um fator de risco comum, o que poderia ser interpretado como a situação econômica do país ou como o risco sistemático de mercado. O risco idiossincrático representa o risco inerente a cada devedor, uma parte de incerteza que não é possível ser estimada pelos fatores de risco. A sua não inclusão é por uma questão de simplicidade, já que seria necessário estimar mais parâmetros de entrada para o modelo, e não poderiam ter sido estimados com a base de dados disponível no presente trabalho.

O algoritmo foi implementado segundo as equações (3.55) e (3.63) com critério de parada sendo o percentil da distribuição de perdas considerado 0,9999 (99,99%) ou quando o número máximo de $n_{it} = 300.000$ iterações for atingido. O modelo CR+ foi implementado no *software* MATLAB®2016a.

Este capítulo está organizado como segue: a seção 5.1 mostra o impacto da probabilidade de *default* na estimação do modelo, dada a suposição de que devem ser suficientemente pequenas, conforme desenvolvimento do capítulo 3; na seção 5.2, o efeito do número de faixas no modelo é analisada, visto que não há um consenso geral sobre o número ideal a ser adotado; por fim, na seção 5.3, os efeitos da volatilidade da probabilidade de *default* são mostrados na estimação do risco de crédito.

5.1 EFEITO DA PROBABILIDADE DE *DEFAULT*

5.1.1 Estimativa de Perda

Conforme visto na seção 3.3, a distribuição de probabilidade de perdas de cada devedor é modelada por uma distribuição Bernoulli. A partir disso, uma das principais suposições no desenvolvimento do modelo CreditRisk+ é a probabilidade de *default* dos devedores ser suficientemente pequena. Essa suposição é necessária pois possibilita a aproximação da distribuição de Bernoulli por uma distribuição de Poisson, sendo então obtida por meio da aproximação de $\ln(1 + w)$ na expansão de Taylor, conforme feito em (3.14),

$$\ln(1 + w) \approx w$$

em que $w = p_A^S(z^{v_A} - 1)$. Sendo assim, é essencial para o desenvolvimento posterior do modelo CR+ e a construção da sua forma analítica.

A Tabela 15 apresenta uma simulação dos valores reais da aproximação realizada e $\ln(1 + w)$, a fim de comparar o comportamento da aproximação.

TABELA 15 – Aproximação logarítmica pela expansão de Taylor

w	$\ln(1 + w)$	Erro (%)
0,02	0,020	1,00
0,05	0,049	2,48
0,10	0,0950	4,92
0,20	0,182	9,70
0,30	0,262	14,34
0,40	0,336	18,88
0,50	0,405	23,32
0,60	0,470	27,66
0,70	0,531	31,92
0,80	0,588	36,10
0,90	0,642	40,22
1,00	0,693	44,27

Conforme o valor da expressão $\ln(1 + w)$ cresce, a aproximação linear w perde precisão, resultando em diferenças cada vez maiores. Dessa forma, quanto maior o valor da probabilidade de *default* p_A , maior será o valor de w , e o erro na aproximação faz com que a estimação seja feita com valores inflacionados. Assim, a estimação da perda da carteira pode ser comprometida por probabilidades de *default* altas.

Segundo Gundlach e Lehrbass (2004), devedores com probabilidade de *default* acima de 0,09 já oferecem uma aproximação imprecisa na estimação. O que pode levar ao modelo superestimar o risco de crédito, e em certos casos, fazer com que o risco estimado seja maior que a soma das exposições líquidas. Neste caso, poderia se entender que existe um risco de perder mais capital do que foi emprestado ou financiado aos devedores.

No entanto, conforme a estimação da probabilidade de *default* realizada no capítulo anterior pela técnica de regressão logística, nem todas as probabilidades de *default* calculadas nas duas carteiras de crédito satisfazem tal condição, tal como as Tabelas 11 e 14 mostram. Na carteira 1, 217 devedores possuem probabilidade de entrar em inadimplência menor que 0,09, o que representa 42,13% do total de devedores. Na segunda carteira de crédito de empréstimos, tal proporção é menor, com 227 dos 1000 devedores tendo uma probabilidade de *default* menor que este limite.

Uma alternativa para as imprecisões causadas por taxas de *default* maiores seria considerar todos os devedores com probabilidade de *default* acima de certo limite como inadimplentes. Dessa forma, define-se um certo valor de corte da probabilidade de *default*, os devedores acima deste limite são considerados como perdas “determinísticas” iguais às suas perdas esperadas e não são considerados na estimação do modelo CR+. Esta metodologia foi utilizada por Schechtman et al. (2004) na estimação do risco de crédito de bancos brasileiros, entretanto esse trabalho utiliza um limite de 0,15 para a probabilidade de *default*.

Inicialmente estima-se o risco de crédito para o modelo padrão, ou seja, sob as condições determinadas da introdução deste capítulo (100 faixas de exposição, volatilidade $\omega = 0.5$ e um único setor), e considerando todos os 515 devedores da carteira 1, independente da probabilidade de *default*.

Na Tabela 16, duas características gerais da carteira de crédito em questão são apresentadas. A primeira diz respeito à exposição total da carteira, sendo a soma das exposições \tilde{v}_A individuais de cada devedor que a compõem. A segunda característica é a perda esperada da carteira, sendo a soma das PE_A individuais dos devedores, dadas conforme a expressão (3.1).

TABELA 16 – Carteira 1 - Dados da carteira de crédito

Exposição Total (R\$)	Perda Esperada (R\$)
9.847.473,23	2.991.633,87

A Tabela 17 apresenta os resultados da estimação do risco pelo modelo CR+ para diferentes níveis de confiança. A interpretação que pode ser feita do VaR é que existe uma chance de 10% da perda ser maior que R\$ 5.023.219,83, 5% de chance da perda ser maior que R\$ 5.836.455,14 e apenas 1% de chance da perda ser maior que R\$ 7.573.382,25, considerando o período de um ano. Já a perda dada pelo CVaR indica a perda média incorrida na carteira no caso extremo da perda ser superior ao valor monetário dado pelo VaR.

Conforme o nível de confiança ϵ é incrementado, o valor em perda também sofre um aumento, refletindo o maior percentual adotado para a estimação. Para $\epsilon = 99\%$ o CVaR estimado de R\$ 9.798.359,92 é bem próximo à exposição total da carteira, o que

TABELA 17 – Carteira 1 - Perda estimada da carteira

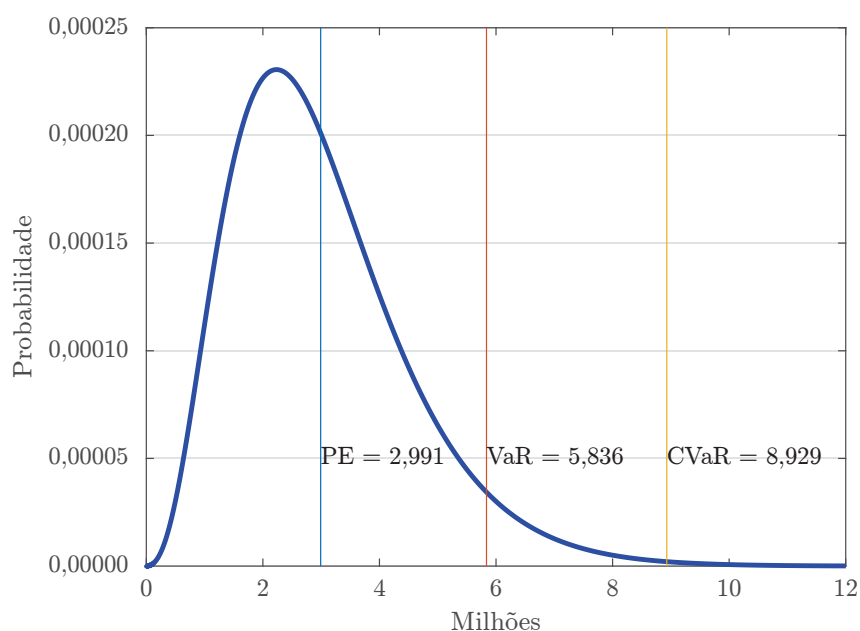
ϵ	VaR (R\$)	CVaR (R\$)
90%	5.023.219,83	8.523.449,91
95%	5.836.455,14	8.929.829,96
99%	7.573.382,25	9.798.359,92

significa que, no caso de 1% da perda extrema ser acima do VaR, existe uma possibilidade de perda esperada de praticamente toda a carteira.

A Figura 5 apresenta a distribuição de probabilidade de perdas estimada para a carteira 1 e as diferentes medidas de risco. A curva foi construída utilizando as probabilidades de perda obtidas pelo algoritmo recursivo do modelo CR+. Esta curva é construída considerando uma distribuição discreta (Poisson), apesar da Figura 5 mostrar uma curva aparentemente contínua, sendo esta continuidade causada pela grande quantidade de capital considerado (na casa dos milhões).

O nível de confiança considerado para o VaR na figura é de 95%, onde as diferentes marcações indicam os valores do VaR e CVaR. Observa-se que a cauda da distribuição de perdas se estende além do valor da exposição total da carteira de crédito no valor de R\$ 9.847.473,23, o que seria consequência das aproximações realizadas no modelo e da probabilidade de *default* utilizada.

FIGURA 5 – Carteira 1 - Distribuição de perdas - modelo padrão



Fonte: A autora (2018)

É possível constatar também que a probabilidade de ocorrência de uma perda de R\$ 2.991.633,87 (o valor da perda esperada) é de aproximadamente 0,020% (ou 0,00020). A perda máxima em um ano, sob um nível de confiança de 95%, é R\$ 5.836.455,14 e no

caso extremo de a perda ser superior a este valor, a perda média incorrida será de R\$ 8.929.829,96.

Para a segunda carteira de crédito estudada neste trabalho, a Tabela 18 apresenta a exposição total e a perda esperada da carteira de crédito. A exposição total é menor em comparação à carteira 1, assim como a perda esperada, sendo aproximadamente três vezes menor em valor monetário.

TABELA 18 – Carteira 2 - Dados da carteira de crédito

Exposição Total (R\$)	Perda Esperada (R\$)
3.271.258,00	1.004.258,26

Seguindo os mesmos parâmetros utilizados no modelo padrão, o risco de crédito da carteira 2 é estimada de forma análoga aos resultados da carteira apresentada anteriormente. Assim, a Tabela 19 apresenta os resultados da estimação do modelo CR+.

TABELA 19 – Carteira 2 - Perda estimada da carteira

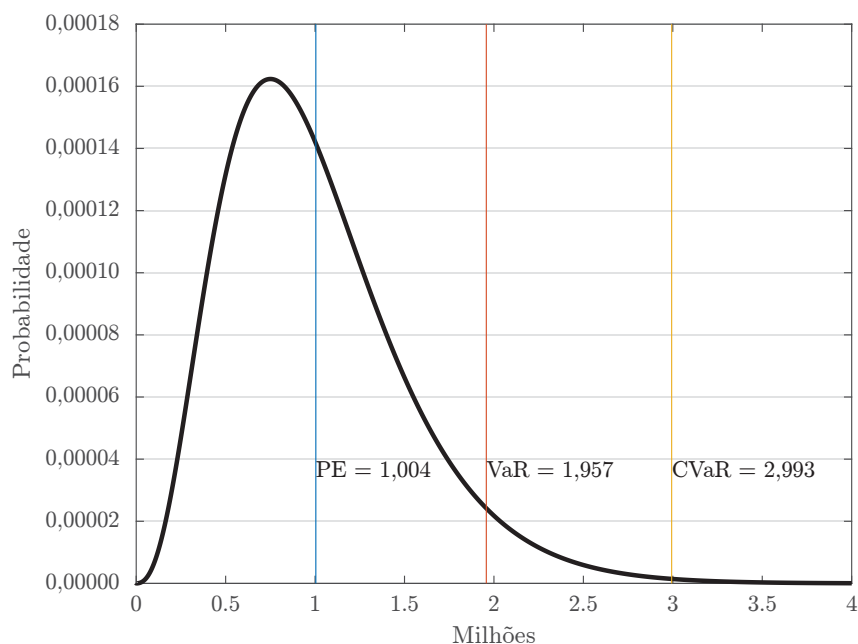
ϵ	VaR (R\$)	CVaR (R\$)
90%	1.684.878,16	2.857.336,00
95%	1.957.315,08	2.993.587,99
99%	2.539.193,15	3.284.492,00

Para o nível de confiança de $\epsilon = 99\%$, temos que a medida de risco CVaR já apresenta valor acima da exposição total da carteira, com perda média no valor de R\$ 3.284.492,00. Este resultado é diferente do que ocorreu na estimação da primeira carteira, o que poderia ser explicado pelas probabilidades de *default* mais altas presentes na carteira atual. De fato, conforme discutido anteriormente, enquanto a carteira 1 possui 42,13% dos devedores com probabilidade de *default* menores que 0,09, na carteira 2 esta proporção é de 22,7%, fazendo com que o risco estimado atualmente seja superestimado.

A Figura 6 mostra a distribuição de probabilidades de perda agregada da carteira. Assim como indicações dos valores da perda esperada, VaR e CVaR. Neste caso, tem-se que a perda máxima da carteira é de R\$ 1.957.315,08, considerando um nível de confiança de 95%. No cenário extremo em que a perda ultrapasse o limite dado pelo VaR, a perda média esperada incorrida deverá ser de aproximadamente R\$ 2.993.587,99. Percebe-se também que a cauda da distribuição de probabilidades ultrapassa o valor total em exposição da carteira, sendo prolongada até 4 milhões de reais, enquanto a exposição da carteira é de aproximadamente 3 milhões.

Dada a suposição de que as probabilidades de *default* precisam ser suficientemente pequenas, pode-se considerar que o risco de crédito foi superestimado nesta seção. Neste caso, o próximo estudo realizada considera a sugestão feita anteriormente, ao definir um limite de corte para a probabilidade de *default* e estimar o risco apenas com os devedores

FIGURA 6 – Carteira 2 - Distribuição de perdas - modelo padrão



Fonte: A autora (2018)

abaixo deste limite. Este processo garantiria as condições para a aplicação do modelo CR+ sem inflacionar o valor da perda.

Dessa forma, adota-se diferentes limites de probabilidades de *default* para a estimação de perdas. Os devedores acima do limite imposto são considerados inadimplentes, sendo a perda incorrida considerada determinística. Para tanto, considera-se os limites de 0,50 e 0,09. O valor de 0,50 é considerado pois foi utilizado como valor de corte na estimação da probabilidade de *default* por meio da técnica de regressão logística, no capítulo 4. No modelo *logit* construído, os devedores acima de 0,50 são considerados inadimplentes, e abaixo, classificados como adimplentes. Segundo Gundlach e Lehrbass (2004) probabilidades de *default* acima de 0,09 já podem causar imprecisões no modelo, sendo também adotado como um valor de limite.

5.1.2 Probabilidade de *default* abaixo de 50% e 9%

Considerando apenas os devedores que possuem probabilidade de *default* abaixo de 0,50, o número de devedores da carteira passa a ser 338 para a carteira 1. Sendo assim, 177 devedores são considerados inadimplentes e a soma de suas perdas esperadas é considerada como uma perda realizada, totalizando R\$ 2.404.564,33. A exposição total e a perda esperada da carteira são apresentadas na Tabela 20.

O risco de crédito é então estimado pelo modelo CR+ seguindo os mesmos parâmetros do modelo padrão. A Tabela 21 apresenta a perda estimada para a carteira.

Como os devedores com $PD > 0,5$ não são considerados nesta análise, é notável

TABELA 20 – Carteira 1 - Dados da carteira de crédito para $PD < 0,5$

Exposição Total (R\$)	Perda Esperada (R\$)
6.618.347,88	587.069,53

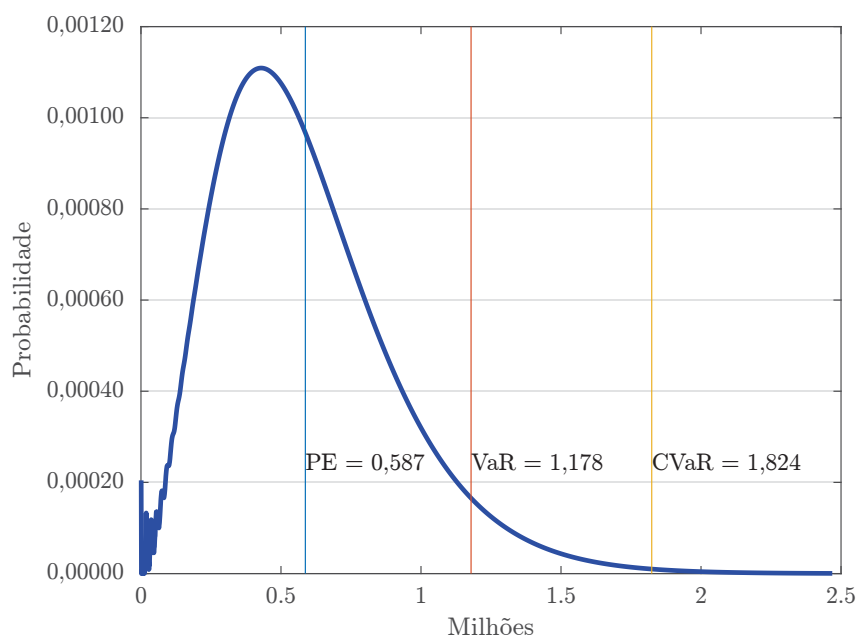
TABELA 21 – Carteira 1 - Perda estimada da carteira para $PD < 0,5$

ϵ	VaR (R\$)	CVaR (R\$)
90%	1.009.489,66	1.739.399,67
95%	1.178.873,32	1.824.029,70
99%	1.540.696,43	2.004.989,50

que os valores estimados de perda da carteira sofrem redução considerável em relação à carteira original. Apesar da exposição total da carteira atual ser naturalmente menor, a relação entre a exposição total e a perda esperada é modificada. Na carteira original, como pode ser visto na Tabela 16, a exposição total da carteira é aproximadamente três vezes maior que o valor da perda esperada, enquanto que na carteira em questão, a exposição de R\$ 6.618.347,88 é onze vezes maior que a perda esperada de R\$ 587.069,53.

Os valores de perda estimados também refletem essa mudança na probabilidade de *default* dos devedores. O VaR e CVaR estimados na Tabela 21 são condizentes com o valor da exposição da carteira, no sentido de que a perda calculada não ultrapassa o valor de exposição total.

Na Figura 7, tem-se a distribuição de probabilidades de perda para o modelo atual e os valores do VaR e CVaR para o nível de confiança de 95%.

FIGURA 7 – Carteira 1 - Distribuição de perdas para $PD < 0,5$ 

Fonte: A autora (2018)

Para o modelo padrão com a carteira original, visto na Figura 5, apesar da exposição total da carteira ser de aproximadamente 10 milhões de reais, a cauda da distribuição se estendia até 12 milhões, sendo considerado um resultado inconsistente com a teoria, indicando que seria possível perder mais do que havia sido emprestado. Apesar da presente distribuição ser semelhante ao do modelo original, percebe-se que a cauda da distribuição na Figura 7 se prolonga até aproximadamente 2,5 milhões de reais, sendo consideravelmente menor que o valor de exposição de 6 milhões. Constata-se, portanto, que probabilidades de *default* altas na carteira podem causar imprecisões na estimação do risco de crédito da carteira.

Considerando o limite da probabilidade de *default* de 0,09 para a carteira 1, mais devedores são retirados da estimação do risco e a carteira conta com 217 devedores dos 515 da carteira completa. Dessa forma, o valor em exposição e a perda esperada desta carteira podem ser vistos na Tabela 22.

TABELA 22 – Carteira 1 - Dados da carteira de crédito para $PD < 0,09$

Exposição Total (R\$)	Perda Esperada (R\$)
4.328.278,32	55.105,53

A estimação do VaR e CVaR sob as atuais condições são vistas na Tabela 23.

TABELA 23 – Carteira 1 - Perda estimada da carteira para $PD < 0,09$

ϵ	VaR (R\$)	CVaR (R\$)
90%	123.037,60	252.484,24
95%	152.540,55	267.019,30
99%	215.977,89	298.790,60

As tabelas anteriores mostram resultados semelhantes aos encontrados quando a carteira desconsidera os devedores com probabilidade de *default* maiores que 0,50. Da mesma forma, a perda esperada da carteira e as medidas de risco sofrem redução considerável.

Dada a suposição teórica do modelo CreditRisk+ em relação à probabilidade de *default*, o risco de crédito estimado sob as condições da seção atual são válidas, especialmente para o modelo que desconsidera devedores com probabilidade de *default* acima de 0,09. Cabe ressaltar que o risco de crédito dos devedores restantes é avaliado por meio do cálculo de sua perda esperada sendo, para o último modelo estimado, no valor de R\$ 2.936.528,34. Este valor de risco é considerado como perda determinística, e portanto, considera-se que a carteira terá uma perda deste valor monetário no período de estimação de um ano.

O estudo atual foi realizada com o intuito de expor a estimação do risco de crédito por meio do modelo CreditRisk+ sob as condições ideais de probabilidade de *default*. Se

a carteira de crédito da empresa de securitização fosse composta por apenas estes 217 devedores, os valores estimados de risco seriam apenas os apresentados na Tabela 23.

A mesma análise pode ser feita para a carteira de crédito alemã. Considerando apenas os clientes com probabilidade de *default* acima do valor de corte de 0,5, a carteira resultante possui 763 devedores, e a exposição total e perda esperada podem ser vistas na Tabela 24.

TABELA 24 – Carteira 2 - Dados da carteira de crédito para $PD < 0,5$

Exposição Total (R\$)	Perda Esperada (R\$)
2.501.005,00	478.160,76

O risco de crédito é então estimado pelo modelo CR+. A Tabela 25 apresenta a perda estimada para a carteira 2.

TABELA 25 – Carteira 2 - Perda estimada da carteira para $PD < 0,5$

ϵ	VaR (R\$)	CVaR (R\$)
90%	806.281,60	1.371.719,99
95%	937.656,97	1.437.407,99
99%	1.218.256,84	1.577.707,99

A perda esperada dos devedores que não foram incluídos na estimação dos riscos de crédito é dada por R\$ 526.097,50.

Para o modelo considerando apenas os devedores que possuem probabilidade de *default* abaixo de 0,09, tem-se uma carteira de crédito composta por 227 devedores, uma redução considerável em comparação aos 1000 devedores na carteira original. As informações básicas desta carteira são apresentadas na Tabela 26.

TABELA 26 – Carteira 2 - Dados da carteira de crédito para $PD < 0,09$

Exposição Total (R\$)	Perda Esperada (R\$)
754.434,00	36.080,62

O risco de crédito dos devedores que foram desconsiderados na estimação atual pode ser representado pela perda esperada dos mesmos, somando um valor de R\$ 968.177,64. O risco medido pelas métricas VaR e CVaR da carteira 2 sob essas condições pode ser visto na Tabela 27.

TABELA 27 – Carteira 2 - Perda estimada da carteira para $PD < 0,09$

ϵ	VaR (R\$)	CVaR (R\$)
90%	67.056,05	121.642,33
95%	79.663,05	127.889,98
99%	106.645,38	141.413,83

Portanto, tem-se que a probabilidade de *default* dos devedores da carteira influencia na estimação do risco de crédito do modelo. Quando estas não são suficientemente pequenas, podem provocar imprecisões no cálculo da distribuição de perdas, o que consequentemente causa uma superestimação dos valores do VaR e CVaR.

Uma das alternativas adotadas para aplicação do modelo CR+, apesar da carteira não possuir probabilidades pequenas o suficiente, é omitir os devedores que possuem maior chance de entrarem em inadimplência. Desta forma, é possível quantificar o risco de crédito sob suposições válidas do CR+.

5.2 EFEITO DO NÚMERO DE FAIXAS E TAMANHO DA UNIDADE DE PERDAS

Uma das características do modelo CreditRisk+ que contribui para sua eficiência computacional é utilizar valores discretos. Isto é, as exposições individuais dos devedores, inicialmente valores contínuos, são ajustadas para números inteiros e agregadas por faixas de exposição. Este processo limita a quantidade de valores de exposição, o que reduz o tempo de computação da distribuição de perdas.

A primeira etapa deste processo é, geralmente, adotar um número de faixas F , e o tamanho da faixa U (amplitude da faixa, ou unidade de perda) é calculado, como desenvolvido em (3.3)

$$U = \frac{\max[\tilde{v}_A]}{F}.$$

Após o cálculo de U , as exposições individuais dos devedores também são ajustadas, conforme (3.2):

$$v_A = \left\lceil \frac{\tilde{v}_A}{U} \right\rceil.$$

Finalmente, as exposições v_A são então agregadas em faixas de acordo com o seu valor. Sendo f uma faixa de exposição, com $f = 1, \dots, F$, um devedor A será atribuído à faixa f se $v_A = f$. Deste modo, em vez de trabalhar com N exposições, a estimação do modelo é feita considerando F faixas de exposições. Da mesma forma, os valores das exposições originais \tilde{v}_A são reduzidos consideravelmente durante o processo de ajuste. Originalmente uma exposição \tilde{v}_A que poderia variar entre centenas ou milhões de reais é ajustada para um valor discreto v_A , que assume valores entre 1 e F . Este processo reduz o custo computacional quando o número de faixas adotado é menor que o número de devedores presentes na carteira.

O propósito de trabalhar com exposições discretas no modelo CR+ é oferecer uma maior eficiência computacional ao algoritmo por meio desta aproximação e ajuste. Deste modo, quanto menor o número de faixas menos processamento é necessário e quanto maior o número de faixas mais próximo se está da exposição original. Dada a natureza do processo de ajuste de exposições, o maior número de faixas possível é $\max[\tilde{v}_A]$ e, neste

caso, a unidade de perdas seria $U = 1$. Neste contexto, um número muito alto tornaria a estimação computacionalmente inviável. Ao mesmo tempo, um número baixo de faixas comprometeria a estimação da distribuição, pois o número de aproximações realizadas ao ajustar as exposições individuais é maior.

À vista disso, a quantidade de faixas F pode ter impactos da estimação do modelo. Então, de forma a analisar o efeito de diferentes números de faixas, considera-se a estimação do modelo CR+ para 50, 300 e 500 faixas, além do valor utilizado no modelo padrão de 100. Outros valores poderiam ser utilizados para esta análise e não há um critério teórico para a escolha dos números de faixas de exposição considerados.

As Tabelas 28 e 29 mostram a estimação do risco de crédito para os diferentes valores de F nas duas medidas de risco utilizadas, respectivamente, o VaR e CVaR.

TABELA 28 – Carteira 1 - Perda estimada (VaR) para diferentes números de faixas

ϵ	Perda da carteira - VaR (R\$)			
	$F = 50$	$F = 100$	$F = 300$	$F = 500$
90%	5.023.334,66	5.023.219,83	5.023.134,60	5.023.113,37
95%	5.836.772,86	5.836.455,14	5.836.231,28	5.836.180,67
99%	7.574.133,42	7.573.382,25	7.572.862,31	7.572.748,91

TABELA 29 – Carteira 1 - Perda estimada (CVaR) para diferentes números de faixas

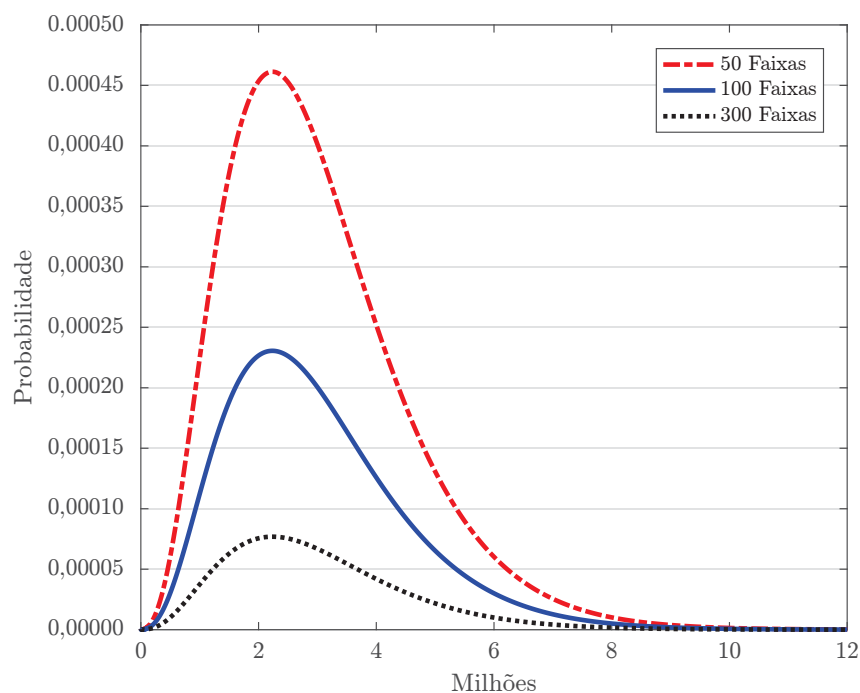
ϵ	Perda da carteira - CVaR (R\$)			
	$F = 50$	$F = 100$	$F = 300$	$F = 500$
90%	8.525.401,16	8.523.449,91	8.522.149,99	8.521.889,99
95%	8.932.041,65	8.929.829,96	8.928.789,98	8.928.425,99
99%	9.800.738,00	9.798.359,92	9.797.059,99	9.796.721,99

Neste caso, conforme o número de faixas adotada aumenta, percebe-se uma diminuição no valor da perda estimada na carteira. Apesar do decréscimo não ser tão significativo, representando alterações na ordem de 0,01% entre um valor de F e o próximo, isto significa que utilizar um número de faixas de exposição pequeno pode acarretar em uma superestimação do risco de crédito.

A Figuras 8 e 9 mostram as distribuições de probabilidade estimadas para diferentes números de faixas. O comportamento da distribuição de probabilidades para diferentes números de faixas varia. Quanto menor o número de faixas utilizadas, como no caso de $F = 50$ na Figura 8, as probabilidades de ocorrência de certos valores de perda são maiores em comparação com $F = 100$.

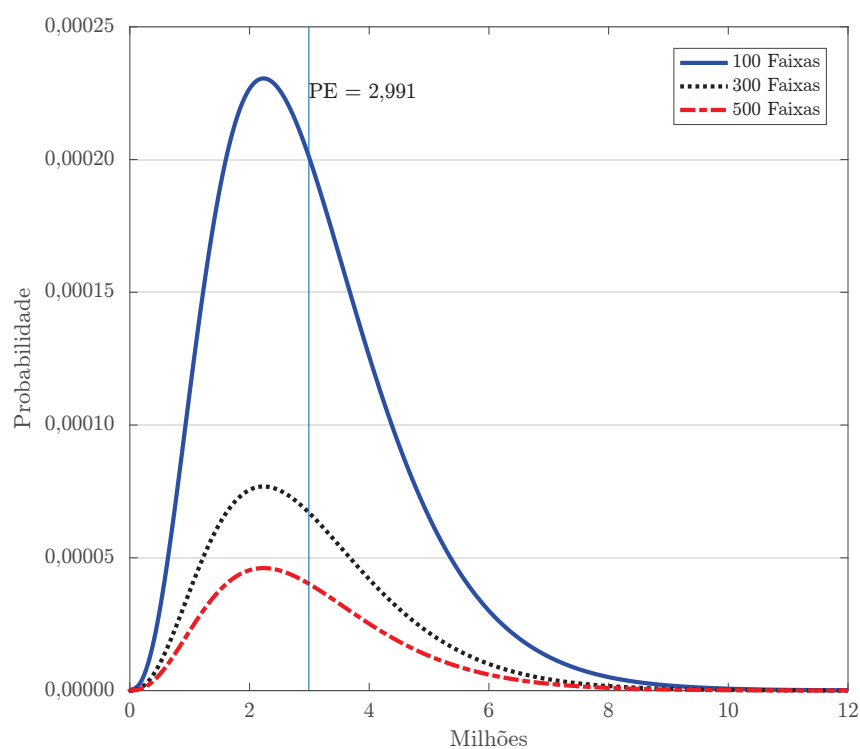
Observando a Figura 9, tem-se o indicativo do valor da perda esperada PE, e para diferentes números de F a probabilidade de perda também varia. Para o modelo padrão, a probabilidade de uma perda agregada do valor da PE é aproximadamente 0,020% (0,00020). Para o mesmo valor de perda no modelo com $F = 300$, a probabilidade é de

FIGURA 8 – Carteira 1 - Distribuição de perdas para diferentes faixas de exposição (50, 100 e 300)



Fonte: A autora (2018)

FIGURA 9 – Carteira 1 - Distribuição de perdas para diferentes faixas de exposição (100, 300 e 500)



Fonte: A autora (2018)

0,0067% (0,000067). Isto pode ser explicado pelo fato de que quanto maior o número de faixas adotado em um modelo, menor a quantidade de devedores que irá compor cada faixa de exposição. Como a probabilidade de *default* de cada faixa depende da probabilidade de *default* dos devedores que estão nesta faixa, decorre que a probabilidade de perda será menor em faixas com menos devedores. Dessa forma, quando maior o número de faixas, menor é a probabilidade de perda por faixa de exposição, o que acarreta nas distribuições de probabilidade de perda terem comportamentos diferentes.

Assim, constata-se que o maior impacto do parâmetro de número de faixas pode ser visto nas Figuras 8 e 9, em relação à curva da distribuição de probabilidades de perdas dos modelos. Dado que, apesar de F influir no valor estimado do risco da carteira, as diferenças no VaR e CVaR não são tão significativas conforme o número de faixas aumenta entre os modelos.

Para a segunda carteira considerada no presente trabalho, os resultados concordam com o que foi encontrado anteriormente. Para esta carteira, além de estimar o risco para os modelos com o mesmo número de faixas como feito na análise da carteira 1, inclui-se a estimação de um modelo com $F = 1000$. As estimações do risco para as medidas de VaR e CVaR podem ser vistas nas tabelas 30 e 31.

TABELA 30 – Carteira 2 - Perda estimada (VaR) para diferentes números de faixas

ϵ	Perda da carteira - VaR (R\$)				
	$F = 50$	$F = 100$	$F = 300$	$F = 500$	$F = 1000$
90%	1.684.915,85	1.684.878,16	1.684.861,60	1.684.857,41	1.684.853,93
95%	1.957.404,75	1.957.315,07	1.957.267,19	1.957.256,32	1.957.247,83
99%	2.539.394,03	2.539.193,15	2.539.078,36	2.539.053,20	2.539.034,02

TABELA 31 – Carteira 2 - Perda estimada (CVaR) para diferentes números de faixas

ϵ	Perda da carteira - CVaR (R\$)				
	$F = 50$	$F = 100$	$F = 300$	$F = 500$	$F = 1000$
90%	2.857.703,97	2.857.335,99	2.857.056,99	2.857.013,99	2.856.968,99
95%	2.994.047,93	2.993.587,99	2.993.269,99	2.993.219,99	2.993.165,99
99%	3.284.951,95	3.284.491,99	3.284.178,99	3.284.117,99	3.284.063,99

Semelhantes aos resultados da carteira 1, observa-se a diminuição do risco conforme a quantidade de faixas é aumentada entre os diferentes modelos. Para as carteiras em análise essa diferença não foi tão significativa na estimação das medidas de risco.

5.3 EFEITO DA VOLATILIDADE DA TAXA DE *DEFAULT*

Conforme discutido na seção 4.4, outro parâmetro importante na estimação do modelo CR+ compreende o σ_k^2 , conhecido na literatura como volatilidade da taxa de *default*.

Este parâmetro insere no modelo a incerteza associada à variações da probabilidade de *default* de cada setor.

No documento original do modelo (CSFP, 1997), uma das sugestões é considerar uma relação linear entre a probabilidade de *default* e a sua volatilidade, definindo σ_k^2 como 50% da p_A de um devedor A , sendo portanto, a volatilidade adotada para o modelo padrão nas investigações feitas até o momento. Contudo, diferentes estudos aplicados a carteiras particulares utilizam outros valores para este parâmetro.

No trabalho de Balzarotti, Falkenheim e Powell (2002), os autores adotaram $\omega = 0.3$ e $\omega = 0.5$ de acordo com dados externos. Em Schechtman et al. (2004), são testados para a taxa de volatilidade entre 20% e 130%. Dessa forma, diferentes proporções foram utilizadas para a definição da volatilidade da probabilidade de *default* e utilizadas para estimar o risco da carteira no presente trabalho. Mais precisamente, considera-se os valores de ω de 0,2, 0,4, 0,6, e 0,8, além do valor de referência 0,5 utilizado nas estimações das seções anteriores.

As Tabelas 32 e 33 mostram as estimações do risco de crédito para os modelos com diferentes volatilidades da taxa de *default*. Conforme o parâmetro ω é acrescido, a perda agregada das carteiras também sofre um aumento considerável a cada mudança do parâmetro.

TABELA 32 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω

ϵ	Perda da carteira - VaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	3.843.270,74	4.623.980,86	5.023.219,83	5.419.897,03	6.188.739,62
95%	4.124.096,21	5.238.253,23	5.836.455,14	6.451.636,66	7.712.027,92
99%	4.684.123,10	6.523.417,03	7.573.382,25	8.697.425,48	11.142.696,63

TABELA 33 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω

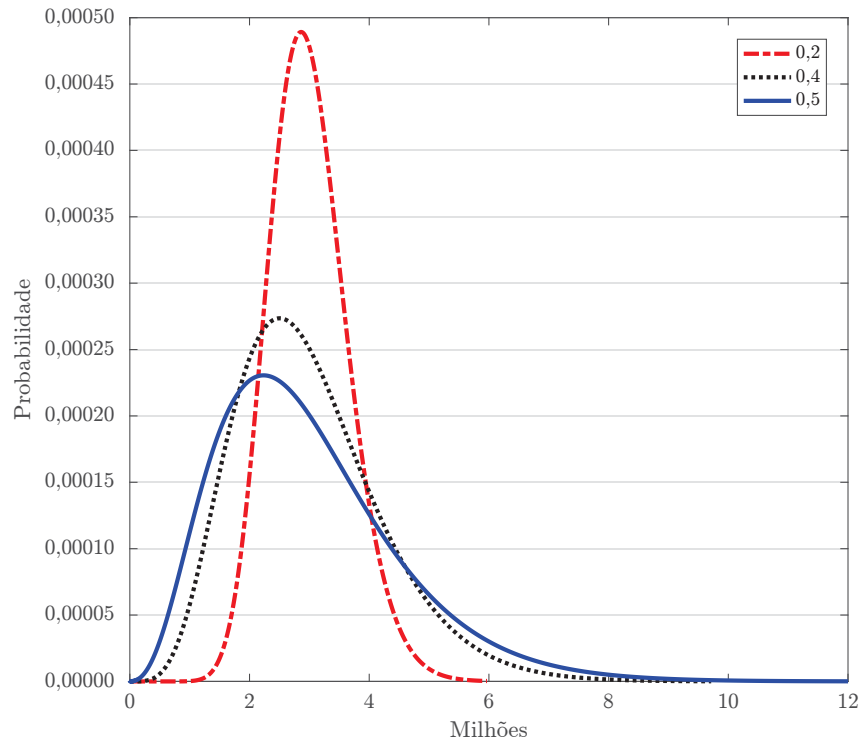
ϵ	Perda da carteira - CVaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	4.906.199,79	7.169.369,90	8.523.449,91	10.015.589,97	13.394.549,97
95%	5.046.599,77	7.476.299,96	8.929.829,96	10.531.559,95	14.156.219,96
99%	5.326.619,66	8.119.019,88	9.798.359,92	11.654.369,95	15.871.439,97

Para a estimação do VaR e do CVaR, observa-se que os modelos com volatilidades de $\omega = 0,6$ e $\omega = 0,8$ produzem estimativas de perda acima do valor total em exposição da carteira. Como citado anteriormente, é uma das fragilidades do modelo CR+ e de sua suposição de que as probabilidades de *default* precisam ser pequenas.

A representação gráfica das distribuições de probabilidade para estes modelos podem ser vistas na Figura 10, que apresenta as curvas para os menores valores de volatilidade, e na Figura 11, que mostra as distribuições dos modelos que utilizaram as

volatilidades de $\omega = 0,5$, $\omega = 0,6$ e $\omega = 0,8$. Dessa forma, a volatilidade da taxa de *default* também altera consideravelmente o comportamento da distribuição de perdas dos modelos.

FIGURA 10 – Carteira 1 - Distribuição de perdas para diferentes volatilidades da PD (0,2, 0,4 e 0,5)



Fonte: A autora (2018)

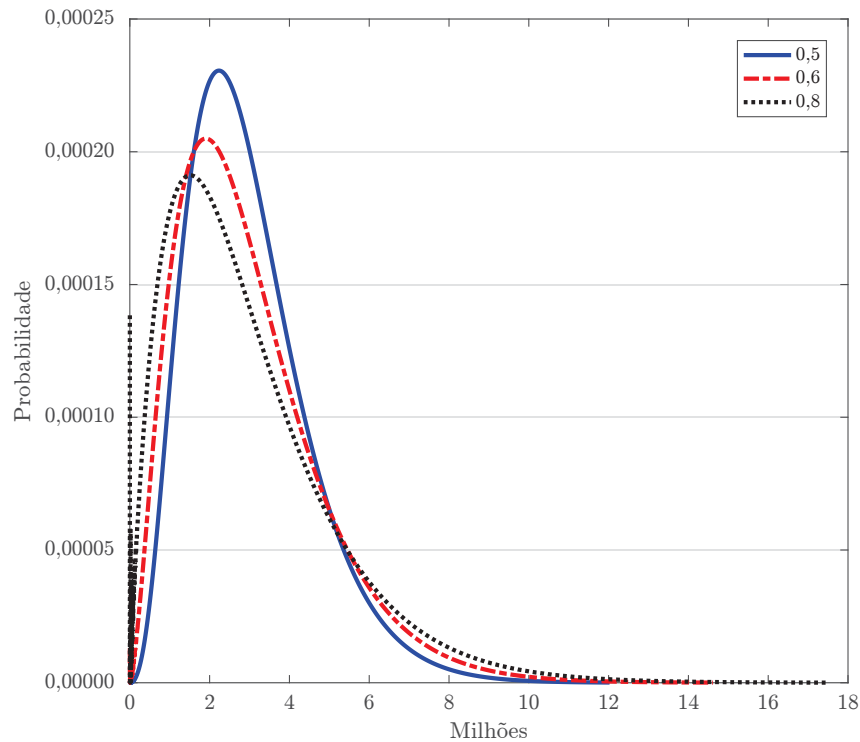
Conforme os modelos utilizam maiores valores do parâmetro ω , observa-se um prolongamento da cauda da distribuição de probabilidades. Para $\omega = 0,2$, observa-se que a probabilidade de perda se estende até aproximadamente 6 milhões de reais na Figura 10. Para $\omega = 0,8$, a cauda da distribuição alcança pouco mais de 17 milhões.

5.3.1 Probabilidade de *default* abaixo de 50% e 9%

Em razão do efeito da probabilidade de *default* na estimação do risco de perda agregada no modelo CR+, esta seção avalia os efeitos da volatilidade das taxas de *default* nos cenários em que se separa a carteira entre os devedores com probabilidades de inadimplência altas ou baixas. O objetivo é observar como o parâmetro ω influencia na estimação da perda quando a probabilidade de *default* é suficientemente pequena de acordo com a teoria do CR+.

Considerando o limite de 0,50 para a probabilidade de *default* para a carteira 1, a quantificação do risco de perda agregada para os diferentes valores de ω é apresentada nas Tabelas 34 e 35.

FIGURA 11 – Carteira 1- Distribuição de perdas para diferentes volatilidades da PD (0,5, 0,6 e 0,8)



Fonte: A autora (2018)

TABELA 34 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω e $PD < 0,05$

ϵ	Perda da carteira - VaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	803.881,34	936.448,04	1.009.489,66	1.083.558,70	1.229.409,36
95%	876.312,36	1.068.300,95	1.178.873,32	1.294.756,19	1.535.666,78
99%	1.021.271,35	1.344.261,10	1.540.696,43	1.754.495,65	2.225.409,36

TABELA 35 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω e $PD < 0,05$

ϵ	Perda da carteira - CVaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	1.080.299,57	1.483.559,76	1.739.399,67	2.025.269,73	2.678.909,82
95%	1.116.569,33	1.549.469,75	1.824.029,70	2.130.569,97	2.831.789,95
99%	1.189.108,77	1.687.529,47	2.004.989,50	2.360.669,67	3.176.939,70

Os resultados são semelhantes ao encontrado para a carteira original, em que volatilidade da probabilidade de *default* faz com que o VaR e CVaR da carteira seja cada vez maior conforme o parâmetro ω cresce. Entretanto, para todos os níveis de confiança ϵ e para todos os ω utilizados, nenhuma estimac o de risco ultrapassa o valor em exposic o total da carteira de R\$ 6.618.347,88.

Para os devedores com probabilidade de *default* abaixo de 0,09, tem-se as estimativas de risco nas Tabelas 36 e 37. Da mesma forma, a perda estimada   menor que o

valor total em exposição da carteiras sob essas condições.

TABELA 36 – Carteira 1 - Perda (VaR) para diferentes volatilidades ω e $PD < 0,09$

ϵ	Perda da carteira - VaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	114.028,44	119.682,51	123.037,60	127.150,63	136.254,88
95%	137.554,67	146.452,85	152.540,55	159.568,68	176.035,48
99%	186.510,96	203.554,60	215.977,89	230.626,50	266.026,18

TABELA 37 – Carteira 1 - Perda (CVaR) para diferentes volatilidades ω e $PD < 0,09$

ϵ	Perda da carteira - CVaR (R\$)				
	$\omega = 0,2$	$\omega = 0,4$	$\omega = 0,5$	$\omega = 0,6$	$\omega = 0,8$
90%	210.911,24	234.908,10	252.484,24	273.778,58	325.831,47
95%	222.740,61	248.091,20	267.019,30	290.002,32	345.773,20
99%	247.077,63	276.819,23	298.790,60	325.491,98	390.727,13

Para as carteiras que desconsideram probabilidades de *default* altas, o fator de volatilidade ω também influencia na quantificação do risco. Quanto maior o ω utilizado, mais incerteza existe em relação à probabilidade de *default* dos devedores, e consequentemente, maior é a perda estimada.

6 CONCLUSÃO

Considerando que a gestão e avaliação de riscos é essencial para qualquer empresa ou instituição financeira, o objetivo central deste trabalho foi quantificar o risco de perda de carteiras de crédito. Como discutido anteriormente, dentre os modelos de risco de crédito existentes, o CreditRisk+ é o que se mostra mais adequado ao cenário brasileiro. Do fato de utilizar dados de entrada limitados e acessíveis, ao contrário de outros modelos, que necessitam de dados por vezes não disponíveis no mercado brasileiro, o modelo CR+ foi selecionado para mensurar o risco de perda.

Dessa forma, o presente trabalho implementou o modelo CR+ para a estimação do risco de crédito de duas carteiras distintas. A primeira, proveniente de uma empresa securitizadora, compreende devedores com dívidas de financiamento de imóveis em um condomínio. A segunda, reúne devedores com empréstimos em uma instituição financeira alemã.

Por meio do algoritmo recursivo proposto pela CSFP (1997), é possível obter a distribuição de probabilidade de perdas da carteira, permitindo assim a estimação de medidas de risco como o Valor em Risco e Valor em Risco Condicional, métricas comumente utilizadas no mercado financeiro.

Uma das principais suposições do modelo é que a probabilidade de *default* dos devedores deve ser pequena. Esta suposição é essencial para a construção da fórmula recursiva no modelo CreditRisk+. O algoritmo recursivo, por sua vez, é um dos fatores que contribui para eficiência computacional do modelo de risco de crédito em estudo.

Neste trabalho, as probabilidades de *default* estimadas por meio da técnica de regressão logística não foram suficientemente pequenas como o modelo sugere, o que ocasionou uma superestimação no valor do VaR e CVaR das carteiras nas análises realizadas. Uma alternativa encontrada foi separar as carteiras entre os devedores que possuíam probabilidade de *default* menor que um dado valor de limite, e aplicar o modelo CR+ apenas para estes devedores específicos. Os devedores com taxa de *default* acima deste limite tiveram o risco mensurado por meio de sua perda esperada. Dessa forma, atende-se as suposições teóricas do modelo CR+ na estimação do risco.

Apesar da estimação da perda ter sido válida sob essas condições, do ponto de vista gerencial não é prático separar uma carteira dessa forma se o objetivo for estimar a perda agregada da carteira. Neste sentido, o modelo CR+ possui deficiências ao exigir taxas de *default* pequenas, não sendo ideal para carteiras de crédito com probabilidades mais altas, como o caso do presente trabalho.

Nos estudos realizadas, percebe-se que o número de faixas de exposição também

influencia no resultado final da estimação da distribuição de perdas. Deste modo, quanto menor o número de faixas adotado, maior é o risco estimado pelo modelo. Constata-se então que um número de faixas pequeno aumenta o valor do VaR e CVaR estimado, fazendo com que o risco calculado pelo modelo seja maior do que realmente deve ser para as carteira de crédito. Além disso, o número de faixas adotado influi no comportamento da curva da distribuição de perdas da carteira, como discutido no capítulo anterior.

O último parâmetro analisado compreende a volatilidade da probabilidade de *default*, sendo o fator que insere no modelo a incerteza referente às taxas de inadimplência. Quando este parâmetro não pode ser estimado diretamente para cada devedor ou setor, a sugestão da literatura é que a volatilidade seja incluída no modelo como uma relação linear da probabilidade de *default*. Sendo assim, diferentes valores de proporção foram considerados para realizar as estimações. Percebe-se então que quanto maior o valor deste parâmetro em um modelo, maior é o risco estimado resultante para a carteira.

Apesar de possuir certas fragilidades em suas suposições teóricas, o modelo CreditRisk+ possui diversas vantagens em sua implementação e aplicação. Além de exigir uma quantidade limitada e acessível de dados de entrada, a sua forma analítica e seu algoritmo recursivo contribuem para uma estimação computacionalmente eficiente. Sendo assim ideal para a realização de simulações com diferentes parâmetros e cenários extremos da economia.

Como sugestão para trabalhos futuros, sugere-se investigar extensões do modelo ou alternativas para evitar os erros de aproximação provenientes das probabilidades de *default* altas.

REFERÊNCIAS

- ACERBI, C.; TASCHE, D. On the coherence of expected shortfall. **Journal of Banking & Finance**, v. 26, p. 1487–1503, 2002.
- ANDRADE, F. W. M. **Desenvolvimento de Modelo de Risco de Portfólio para Carteiras de Crédito a Pessoas Físicas**. 2004. Tese (Doutorado em Administração) – Escola de Administração de Empresas de São Paulo, EAESP/FGV, São Paulo.
- ARTZNER, P. et al. Coherent Measures of Risk. **Mathematical Finance**, v. 9, n. 3, p. 203–228, 1999.
- AVELINO, L. F. C. **Aspectos Jurídicos da Securitização no Brasil**. 2014. Dissertação (Mestrado em Direito) – Faculdade de Direito da Universidade de São Paulo, São Paulo.
- AVESANI, R.G. et al. Review and Implementation of Credit Risk Models. In: ONG, L. L. (Org.). **A guide to IMF stress testing : methods and models**. Washington, D.C.: International Monetary Fund, 2014. cap. 10, p. 135–149.
- BALZAROTTI, V.; FALKENHEIM, M.; POWELL, A. On the Use of Portfolio Risk Models and Capital Requirements in Emerging Markets: The Case of Argentina. **The World Bank Economic Review**, v. 16, n. 2, p. 197–212, 2002.
- BIELECKI, T.R.; RUTKOWSKI, M. **Credit Risk: Modeling, Valuation and Hedging**. [S.l.]: Springer Berlin Heidelberg, 2004. (Springer Finance).
- BOLTON, C. **Logistic regression and its application in credit scoring**. 2009. Dissertação (Mestrado) – Universidade de Pretoria, África do Sul.
- BRITO, G.A.S.; ASSAF NETO, A. Modelo de Classificação de Risco de Crédito de Empresas. pt. **Revista Contabilidade & Finanças**, v. 19, p. 18–29, abr. 2008. ISSN 1519-7077.
- BURNHAM, K. P.; ANDERSON, D. R. **Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach**. 2. ed. Nova Iorque: Springer-Verlag, 2002. p. 488.
- CAOQUETTE, J.B.; ALTMAN, E.I.; NARAYANAN, P. **Managing Credit Risk: The Next Great Financial Challenge**. 2. ed. [S.l.]: Wiley, 2008. (Frontiers in Finance Series).
- CREDIT SUISSE FINANCIAL PRODUCTS, CSFP. **Credit Risk+: A Credit Risk Management Framework**. Londres, 1997. p. 72.
- CROUHY, M.; GALAI, D.; MARK, R. **Risk Management**. 1. ed. [S.l.]: McGraw-Hill, 2001.

- CROUHY, M.; GALAI, D.; MARK, R. **The Essentials of Risk Management**. 1. ed. [S.l.]: McGraw-Hill, 2005.
- DHEERU, D.; TANISKIDOU, E. K. **UCI Machine Learning Repository**. [S.l.: s.n.], 2018. Disponível em: <<http://archive.ics.uci.edu/ml>>.
- FABOZZI, F. J. **Introduction to Securitization**. Nova Jersey: Wiley & Sons, 2008. p. 366.
- GIESE, G. Enhancing CreditRisk+. **Risk**, v. 16, n. 4, p. 73–77, 2003.
- GORDY, M. B. Saddlepoint approximation of CreditRisk+. **Journal of Banking & Finance**, v. 26, n. 7, p. 1335–1353, 2002.
- GUNDLACH, V. M.; LEHRBASS, F. **CreditRisk+ in the Banking Industry**. [S.l.]: Springer Berlin Heidelberg, 2004. (Springer Finance).
- GUPTON, G. M.; FINGER, C. C.; BHATIA, M. **Credit Metrics. Technical Report**. Nova Iorque, 1997. p. 193.
- HAIR, J. F. et al. **Multivariate Data Analysis**. 7. ed. [S.l.]: Prentice Hall, 2010. (Always learning).
- HAND, D.; HENLEY, W. Statistical classification Methods in Consumer Credit Scoring: a Review. **Journal of the Royal Statistical Society Series A (Statistics in Society)**, v. 160, n. 3, p. 523–541, 1997.
- HOSMER, D.W.; LEMESHOW, S. **Applied Logistic Regression**. 2. ed. [S.l.]: Wiley, 2000. (Wiley series in probability and statistics).
- JAMES, B. R. **Probabilidade: Um curso em Nível Intermediário**. 2. ed. [S.l.]: IMPA, 1996.
- JOHNSON, N. L.; KOTZ, S.; KEMP, A. W. **Univariate Discrete Distributions**. [S.l.]: Wiley & Sons, 1992. p. 565.
- JORION, P. **Value at Risk: The New Benchmark for Managing Financial Risk**. 2. ed. [S.l.]: McGraw-Hill Education, 2001.
- KLUGE, D.; LEHRBASS, F. B. Basics of Credit Risk+. In: GUNDLACH, M.; LEHRBASS, F. (Org.). **CreditRisk+ in the Banking Industry**. [S.l.]: Springer Berlin Heidelberg, 2004. cap. 2, p. 7–24.
- KMV, Corporation. **Portfolio management of default risk**. São Francisco, 1993.
- MELCHIORI, M. R. CreditRisk+ by Fast Fourier Transform. **YieldCurve**, Santa Fé, Universidad Nacional del Litoral, 2004. Disponível em: <<https://ssrn.com/abstract=1122844>>.
- OLIVEIRA, T. D. **Análises Preditivas para um Problema de Securitização**. 2018. Monografia (Graduação em Matemática Industrial) – Departamento de Matemática, Universidade Federal do Paraná, UFPR, Curitiba.

- PANJER, H. H. Recursive Evaluation of a Family of Compound Distributions. **ASTIN Bulletin**, Cambridge University Press, v. 12, n. 1, p. 22–26, 1981.
- PEIXOTO, C. **O que são Certificados de Recebíveis Imobiliários (CRI)**. 2016. Disponível em: <<https://cpeixejr.wordpress.com/2016/11/29/o-que-sao-certificados-de-recebiveis-imobiliarios-cri/>>. Acesso em: 10 dez. 2018.
- R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>.
- REIß, O. Fourier Inversion Techniques for CreditRisk+. In: GUNDLACH, M.; LEHRBASS, F. (Org.). **CreditRisk+ in the Banking Industry**. [S.l.]: Springer Berlin Heidelberg, 2004. cap. 8, p. 111–128.
- ROCKAFELLAR, R. T.; URYASEV, S. Conditional value-at-risk for general loss distributions. **Journal of Banking & Finance**, n. 26, p. 1443–1471, 2002.
- SAUNDERS, A.; ALLEN, L. **Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms**. 2. ed. Nova Iorque: John Wiley & Sons, 2002. p. 336.
- SCHECHTMAN, R. et al. **Credit Risk Measurement and the Regulation of Bank Capital and Provision Requirements in Brazil - A Corporate Analysis**. [S.l.], 2004.
- SING, T. et al. ROCr: visualizing classifier performance in R. **Bioinformatics**, v. 21, n. 20, p. 7881, 2005. Disponível em: <<http://rocr.bioinf.mpi-sb.mpg.de>>.
- VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. 4. ed. Nova Iorque: Springer, 2002. ISBN 0-387-95457-0. Disponível em: <<http://www.stats.ox.ac.uk/pub/MASS4>>.
- VOJTEK, M.; KOČENDA, E. Credit-Scoring Methods. **Czech Journal of Economics and Finance**, Charles University Prague, Faculty of Social Sciences, v. 56, n. 3-4, p. 152–167, 2006.
- WILSON, T. C. Measuring and Managing Credit Portfolio Risk. In: BOL, G.; NAKHAEIZADEH, G.; VOLLMER, KH. (Org.). **Risk Measurement, Econometrics and Neural Networks. Contributions to Economics**. [S.l.]: Physica, Heidelberg, 1998. p. 259–306.